(19)日本国特許庁（ＪＰ）　(12) **公開特許公報（Ａ）**　(11)特許出願公開番号

特開平6－202817

(43)公開日　平成６年(1994)７月22日

(51)Int.Cl.[5]	識別記号	庁内整理番号	ＦＩ	技術表示箇所
Ｇ０６Ｆ　3/06	３０５　Ｃ	7165－5B		
	３０１　Ｚ	7165－5B		

審査請求　未請求　請求項の数34　（全 33 頁）

(21)出願番号　特願平4－348301

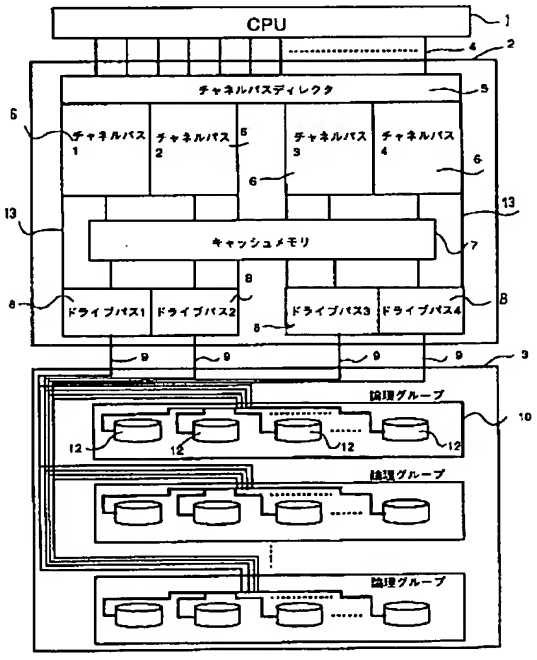
(22)出願日　平成４年(1992)12月28日

(71)出願人　000005108
株式会社日立製作所
東京都千代田区神田駿河台四丁目６番地
(72)発明者　角田　仁
東京都国分寺市東恋ケ窪１丁目280番地
株式会社日立製作所中央研究所内
(74)代理人　弁理士　薄田　利幸

(54)【発明の名称】　ディスクアレイ装置及びそのデータ更新方法

(57)【要約】
【目的】ディスクアレイシステムにおいて、データの書込み時のオーバヘッドを減少させる。
【構成】パリティグループを構成する論理グループ１０に、複数のダミーデータ領域を含ませる。１つの論理グループの各データ及びパリティは、各々別々のＳＣＳＩドライブ１２に分散して格納される。更新時において、更新するデータ及びパリティを書き込む際に、最も効率良く書き込めるＳＣＳＩドライブを判定し、スケジューリングを行なう。更新するデータ及びパリティは更新前のデータ、パリティが格納されていた領域ではなく、このダミーデータの位置に書き込む。
【効果】データの更新時において、更新するデータ及びパリティの書込みをスケジューリングしてダミーデータ領域に効率良く書き込むため、書込み時のオーバヘッドは減少する。また、障害発生時のデータ回復処理や、ドライブ拡張時のデータ書き込み処理が容易となる。

図1



1

【特許請求の範囲】

【請求項１】それぞれ複数のデータと、該複数のデータから生成された少なくとも１つの誤り訂正符号を含む複数のデータグループを生成し、各データグループを構成する複数のデータ及び誤り訂正符号が複数のドライブ内に分散されるようにして保持する記憶装置におけるデータの更新方法であって、

いずれかのデータグループを構成する複数のデータの１つを新たなデータで書き換えるときに、その一つのデータを新たなデータで更新した後の、該データグループに対する新たな誤り訂正符号を新たに生成し、

該新データと該新誤り訂正符号とを、前記記憶装置に含まれた各データグループを分散して書き込むに必要なドライブの数より多いドライブの各々に設けられた予備領域の内、該更新すべきデータグループに対するデータと誤り訂正符号がそれぞれ記憶されているドライブとは異なるドライブに、分散して保持することを特徴とする記憶装置におけるデータの更新方法。

【請求項２】前記書きかえられるべきデータと、該データの属するデーグループの旧誤り訂正符号とを保持していたドライブの領域を新たな予備領域として確保することを特徴とする請求項１記載の記憶装置におけるデータの更新方法。

【請求項３】それぞれ複数のデータと、該複数のデータから生成された少なくとも１つの誤り訂正符号を含む複数のデータグループを生成し、各データグループを構成する複数のデータ及び誤り訂正符号が複数のドライブ内に分散されるようにして保持する記憶装置におけるデータの更新方法であって、

それぞれのデータグループの書き込み時に、それぞれのデータグループに対して、該データグループ内の各データと誤り訂正符号を書き込むドライブと異なるドライブに、少なくとも２個の予備領域を確保し、

いずれかのデータグループに属するデータの一つを新たなデータで更新するときに、該データグループに対して確保された前記少なくとも二つの予備領域の一つに該新たなデータを書き込み、

該新たなデータて更新された後の該データグループに対する新たな誤り訂正符号を生成し、

前記新たな誤り訂正符号を前記少なくとも二つの予備領域の他の１つに書き込み、

前記書き換え前のデータ及び誤り訂正符号がそれぞれ書き込まれていた二つの記憶領域を、上記データグループのための新たな予備領域として保持する、

ことを特徴とする記憶装置におけるデータの更新方法。

【請求項４】前記データグループ内において、前記誤り訂正符号を生成する各データは、別々のドライブの同一の物理的なアドレスに格納し、前記誤り訂正符号も同様に別のドライブの、該誤り訂正符号の生成に関与したデータと同一の物理的なアドレスに格納し、

2

前記予備領域は前記データおよび誤り訂正符号と同様に、前記データ及び誤り訂正符号が格納されているドライブとは別のドライブの同一の物理的なアドレスに各々格納することを特徴とする請求項１または３記載の記憶装置におけるデータの更新方法。

【請求項５】データを格納する複数のドライブ群を有するディスク装置と、外部からのデータの入出力要求に対応して該ディスク装置のデータ格納を管理する制御装置からなるディスクアレイシステムにおけるデータの更新方法であって、該格納されているデータから複数のデータを選択してパリティを生成し、該パリティとその生成に関与したデータの集合をパリティグループの単位とし、各パリティグループのパリティとデータを前記ドライブ群に格納するものにおいて、

外部から１回に読み出しまたは書き込みする単位で転送されてきたデータを前記ドライブ群の各格納領域に分散して格納し、

該分散して格納されているデータから複数のデータを選択してパリティを生成し、該パリティとその生成に関与したデータの集合をパリティグループの単位とし、前記各パリティグループに複数のダミーデータを付与し、

前記各パリティグループにおける、前記複数のダミーデータを、前記パリティとその生成に関与するデータとが格納されているドライブ群とは別のドライブ群に、各々独立して保持することを特徴とするディスクアレイシステムにおけるデータ更新方法。

【請求項６】データを格納する複数のドライブ群を備えたディスク装置と、上位装置からのデータの入出力要求に対応して当該ディスク装置を管理する制御装置からなるディスクアレイシステムにおけるデータ更新方法において、

上記上位装置から１回に読み出しまたは書込みする単位で転送されてきたデータを複数のドライブ群に別々に分散して格納し、

該別々のドライブに格納されているデータから複数のデータを選択し、これらの選択されたデータからパリティを生成し、

この生成したパリティと生成に関与したデータの集合をパリティグループとし、同一のパリティグループにおいては、パリティもデータと同様に、このパリティを生成するのに関与したデータが格納されているのとは別々のドライブに各々格納すると共に、

各パリティグループに対する複数の予備の書込み領域を、パリティを生成するデータと、それらのデータにより生成したパリティが格納されているドライブに各々分散して保持することを特徴とするディスクアレイシステムにおけるデータ更新方法。

【請求項７】予備の書き込み領域を１パリティグループに対し少なくとも２個所以上対応させ、この予備の書き込み領域において、特定のパリティグループに対応した

予備の書き込み領域をダミーデータとし、
前記パリティグループにおいてパリティの生成に関与するデータと、それらのデータより生成したパリティとが格納されているドライブとは別のドライブに、前記それぞれのダミーデータを格納することを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項８】パリティを生成するデータと生成したパリティにより構成されるパリティグループと、予備の書き込み領域が、それぞれ別のドライブに格納されており、これらのドライブの集合を論理グループとし、
該論理グループ内において、パリティを生成する各データは、別々のドライブの同一の物理的なアドレスに格納し、パリティも同様に別のドライブの、パリティの生成に関与したデータと同一の物理的なアドレスに格納し、ダミーデータはデータおよびパリティと同様に、データ、パリティが格納されているドライブとは別のドライブの同一の物理的なアドレスにもつことを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項９】パリティを生成するデータと生成したパリティにより構成されるパリティグループと、予備の書き込み領域が、それぞれ別のドライブに格納されており、これらのドライブの集合を論理グループとし、
前記ディスク装置を制御する制御装置は、上位装置からの書込み要求に対し、論理グループ内において、上位装置が書き込み先として指定してきたデータのアドレスを前記ダミーデータの物理的なアドレスに変換して書き込むことを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項１０】パリティを生成するデータと生成したパリティにより構成されるパリティグループと、予備の書き込み領域が、それぞれ別のドライブに格納されており、これらのドライブの集合を論理グループとし、
前記ディスク装置を制御する制御装置は、上位装置から書込み要求が発行された場合、
上位装置が書込み先に指定したアドレスに対応する物理的なアドレスに書き込まれている旧データが格納されているドライブが属する論理グループにおいて、その論理グループ内の旧データおよび更新される旧パリティおよびダミーデータが格納されている各ドライブに対して、各々の回転位置を検出し、
該検出された回転位置の情報から、回転待ち時間を算出し、
該回転位置または回転待ち時間の情報から、ダミーデータの格納されている複数のドライブ群の中で、どのドライブ内のダミーデータの物理的なアドレスに論理グループの新たなデータを格納するかを判定することを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項１１】前記ディスク装置がキャッシュメモリを備えており、パリティを生成するデータと生成したパリティにより構成されるパリティグループと予備の書き込み領域がそれぞれ別のドライブに格納されており、これらのドライブの集合を論理グループとし、
前記論理グループにおいてダミーデータの格納されている複数のドライブ群の中から、実際に新データを格納するドライブを判定した後、そのドライブに対し先行して書込み要求を発行して新データを書き込み、同時に前記キャッシュメモリ内に新データのコピーを残し、
該キャッシュメモリ内の新データのコピーにより、新パリティの作成を行い、該新パリティの作成完了後、この新パリティを新データを格納したのとは別のダミーデータの物理的なアドレスに書き込むことを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項１２】パリティを生成するデータと生成したパリティにより構成されるパリティグループと、予備の書き込み領域が、それぞれ別のドライブに格納されており、これらのドライブの集合を論理グループとし、
前記論理グループ内のあるドライブに障害が発生したとき、該障害が発生したドライブ内のデータまたはパリティを回復処理により復元し、この復元したデータまたはパリティを予備の書き込み領域に書き込み、
前記障害が発生したドライブを正常なドライブに交換し、
前記交換後、この交換した正常なドライブは全てダミーデータにより構成される予備の書き込み領域により構成されているとして論理グループを再構成して処理を再開することを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項１３】パリティを生成するデータと生成したパリティにより構成されるパリティグループと、予備の書き込み領域が、それぞれ別のドライブに格納されており、これらのドライブの集合を論理グループとし、
前記論理グループ内のあるドライブに障害が発生したことを感知した場合、障害が発生したドライブを正常なドライブに交換し、
前記障害が発生したドライブ内のデータまたはパリティを回復処理により復元し、
この復元したデータまたはパリティと、障害が発生したドライブ内にあったダミーデータを、交換した正常なドライブに書き込んで論理グループを再構成して処理を再開することを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項１４】パリティを生成するデータと生成したパリティにより構成されるパリティグループと、予備の書き込み領域が、それぞれ別のドライブに格納されており、これらのドライブの集合を論理グループとし、
前記論理グループ内のダミーデータの数を、書込み処理

時間の要求にあわせて変更可能としたことを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項１５】パリティを生成するデータと生成したパリティにより構成されるパリティグループと、予備の書き込み領域が、それぞれ別のドライブに格納されており、これらのドライブの集合を論理グループとし、

前記論理グループ内のダミーデータの数を、障害発生時に対する信頼性の要求にあわせて変更可能としたことを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項１６】同一パリティグループに対してて、ダミーデータ数の異なるグループを混在させることを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項１７】前記パリティグループにおけるデータとパリティとダミーデータの構成を、テーブルにより管理することを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項１８】複数パリティグループ間で、予備領域を共有することを特徴とする請求項５記載の記憶装置におけるデータ更新方法。

【請求項１９】パリティを作成するデータの集合と作成されたパリティにより構成されるパリティグループと、ダミーデータでサブ論理グループを構成し、各論理グループをこのサブ論理グループにより構成し、

前記複数の論理グループ間でダミーデータを共有する場合、サブ論理グループのデータとパリティとダミーデータの割り当て方を変えることにより実現することを特徴とする請求項５または６記載のディスクアレイシステムにおけるデータ更新方法。

【請求項２０】前記サブ論理グループを構成するデータとパリティとダミーデータの割り当てを、テーブルにより管理することを特徴とする請求項１９記載のディスクアレイシステムにおけるデータ更新方法。

【請求項２１】前記サブ論理グループを構成するデータとパリティとダミーデータの割り当てを、ユーザが自由に設定することを可能とすることを特徴とする請求項１９記載のディスクアレイシステムにおけるデータ更新方法。

【請求項２２】前記サブ論理グループを構成するデータとパリティとダミーデータの割り当てを、上位装置からの読み出し要求または書込み要求が、競合しないような論理グループにおいて構成することを特徴とする請求項１９記載のディスクアレイシステムにおけるデータ更新方法。

【請求項２３】書き込むべき複数のデータから少なくとも１つの誤り訂正符号を含む１つのデタループを生成し、該データグループを構成する複数のデータ及び誤り訂正符号を複数のドライブ群内に格納する制御装置を有

する記憶装置であって、該制御装置が、

前記データグループを構成する複数のデータの１つを新たなデータに書き換える更新要求に応答して、該更新データと新たな誤り訂正符号とを含む１つのデータグループを新たに生成する手段と、

前記データグループの更新データと誤り訂正符号とを、前記書き換え前のデータおよび誤り訂正符号が格納されていたドライブとは異なるドライブにそれぞれ格納する手段とを備えたことを特徴とする記憶装置。

【請求項２４】書き込むべき複数のデータから少なくとも１つの誤り訂正符号を含む１つのデータグループを生成し、該データグループを構成する複数のデータ及び誤り訂正符号を複数のドライブ群内に格納する制御装置を有する記憶装置であって、該制御装置が、

データの書き込み要求に応答して、前記データグループに対して、該データグループ内の各データと誤り訂正符号及び少なくとも２個の予備領域からなる複数の記憶領域を前記記憶装置の互いに異なるドライブ内に各々確保する手段と、

前記確保された複数の記憶領域に前記データグループ内の各データと誤り訂正符号を各々書き込み、残りの少なくとも二つの記憶領域を前記予備領域として保持する手段と、

記憶された前記データの一つを新たなデータで更新する要求に応答して、前記少なくとも二つの予備領域の一つに該新たなデータを書き込む手段と、

前記新たなデータに更新された後の複数のデータに対する新たな誤り訂正符号を含む新たなデータグループを生成する手段と、

前記新たな誤り訂正符号を前記少なくとも二つの予備領域の他の１つに書き込み、前記複数の記憶領域のうち、書き換えられた前記複数のデータ及び誤り訂正符号が書き込まれていた少なくとも二つの記憶領域を上記更新されたデータグループのための新たな予備領域として保持する手段、

とを備えたことを特徴とする記憶装置。

【請求項２５】データを格納する複数のドライブ群を有するディスク装置と、外部からのデータの入出力要求に対応して該ディスク装置のデータ格納を管理する制御装置からなり、該制御装置が、前記格納されているデータから複数のデータを選択してパリティを生成し、該パリティとその生成に関与したデータの集合をパリティグループの単位とし、各パリティグループのパリティとデータを前記ドライブ群に格納するディスクアレイシステムにおいて、

前記制御装置が、

外部から１回に読み出しまたは書き込みする単位で転送されてきたデータを前記ドライブ群の各格納領域に分散して格納する手段と、

該分散して格納されているデータから複数のデータを選

択してパリティを生成し、該パリティとその生成に関与したデータの集合をパリティグループの単位とし、前記各パリティグループに複数のダミーデータを付与する手段と、

前記各パリティグループにおける、前記複数のダミーデータを、前記パリティとその生成に関与するデータとが格納されているドライブ群とは別のドライブ群に、各々独立して格納する手段とを備えたことを特徴とするディスクアレイシステム。

【請求項２６】データを格納する複数のドライブ群を備えたディスク装置と、上位装置からのデータの入出力要求に対応して当該ディスク装置を管理する制御装置からなるディスクアレイシステムにおいて、前記制御装置が、

上記上位装置から１回に読み出しまたは書込みする単位で転送されてきたデータを複数のドライブ群に別々に分散して格納する手段と、

該別々のドライブに格納されているデータから複数のデータを選択し、これらの選択されたデータからパリティを生成するパリティ生成回路と、

この生成したパリティと生成に関与したデータの集合をパリティグループとし、同一のパリティグループにおいては、パリティもデータと同様に、このパリティを生成するのに関与したデータが格納されているのとは別々のドライブに各々格納すると共に、各パリティグループに対する複数の予備の書込み領域を、パリティを生成するデータと、それらのデータにより生成したパリティが格納されているドライブに各々分散して確保する手段とを備えたことを特徴とするディスクアレイシステム。

【請求項２７】前記制御装置は、上位装置からの書込み要求に対し、パリティグループ内において、上位装置が書き込み先として指定してきたデータのアドレスを前記ダミーデータの物理的なアドレスに変換して書き込むアドレス変換用のテーブルを備えていることを特徴とする請求項２５または２６記載のディスクアレイシステム。

【請求項２８】前記制御装置は、上位装置から書込み要求が発行された場合、上位装置が書込み先に指定したアドレスに対応する物理的なアドレスに書き込まれている当該旧データが格納されているドライブが属する論理グループにおいて、その論理グループ内の当該旧データおよび更新される当該旧パリティ及びダミーデータが格納されている各当該ドライブに対して、各々の回転位置を検出する回転位置検出回路と、

該検出された回転位置の情報から、回転待ち時間を算出する手段と、

該回転位置または回転待ち時間の情報から、ダミーデータの格納されている複数のドライブ群の中で、どのドライブ内のダミーデータの物理的なアドレスに論理グループの新たなデータを格納するかを判定する手段とを備えていることを特徴とする請求項２５または２６記載のディスクアレイシステム。

【請求項２９】前記制御装置は、上位装置から書込み要求が発行された場合、パリティグループにおいて当該旧データと当該旧パリティの格納されているドライブ群の中で、回転待ち時間の大きい方のドライブを基準にし、その、基準のドライブから当該旧データまたは当該旧パリティを読み出して、更新する新しいパリティを作成すると共に、ダミーデータの格納されている複数のドライブ群の中で、一番早くその新パリティを書き込めるドライブを判定する手段とを備えていることを特徴とする請求項２５または２６記載のディスクアレイシステム。

【請求項３０】前記制御装置は、前記パリティグループのドライブが属する論理グループ内のあるドライブに障害が発生したとき、該障害が発生したドライブ内のデータまたはパリティを回復処理により復元する手段と、

該復元データまたはパリティを前記予備の書き込み領域に書き込む手段と、

前記障害が発生したドライブを正常なドライブに交換し、該交換した正常なドライブ群は全てダミーデータにより構成される予備の書き込み領域により構成されているとして論理グループを再構成して処理を再開する手段とを備えていることを特徴とする請求項２５または２６記載のディスクアレイシステム。

【請求項３１】前記制御装置は、前記パリティグループのドライブが属する論理グループ内において、データとパリティとダミーデータの構成を管理するテーブルを備えていることを特徴とする請求項２５または２６記載のディスクアレイシステム。

【請求項３２】前記制御装置は、サブ論理グループを構成するデータとパリティとダミーデータの割り当てを管理するテーブルを備えていることを特徴とする請求項３１記載のディスクアレイシステム。

【請求項３３】前記制御装置は、前記パリティグループのドライブが属する論理グループ内のあるドライブに障害が発生した場合、正常な残りのドライブ内のデータとパリティから、障害が発生したドライブ内のデータまたはパリティを回復処理により復元すると共に、この復元したデータまたはパリティを予備の書き込み領域に書き込む制御を行うプロセッサ持つことを特徴とする請求項２５または２６記載のディスクアレイシステム。

【請求項３４】前記制御装置は、前記パリティグループのドライブが属する論理グループ単位に、その内部にアドレス変換用のテーブルと、回転位置検出回路と、パリティ生成回路と、キャッシュメモリと、それらを制御するマイクロプロセッサを持つことを特徴とする請求項２５または２６記載のディスクアレイシステム。

【発明の詳細な説明】

【０００１】

【産業上の利用分野】本発明はディスクアレイ装置に関するものである。

【０００２】

【従来の技術】近年高度情報化に伴い、コンピュータシステムにおいて、２次記憶装置の高性能化が要求されてきた。その一つの解として、多数の比較的容量の小さなディスク装置（以下ドライブ）により構成されるディスクアレイが考えられている。例えば、次の論文にそのような例が開示されている。

【０００３】「D. Patterson, G. Gibson, and R. H. Kartz ; A Case for Redundant Arrays of Inexpensive Disks(RAID), in ACM SIGMOD Conference, Chicago, IL, (June1988)」

上記論文において、データを分割して並列に処理を行うディスクアレイ（レベル３）とデータを分散して、独立に扱うディスクアレイ（レベル５）について、その性能および信頼性の検討結果が報告されている。

【０００４】以下にデータを分散して、独立に扱うディスクアレイ（レベル５）について説明する。レベル５のディスクアレイでは個々のデータを分割せずに独立に扱い、多数の比較的容量の小さなドライブに分散して格納するものである。現在、一般に使用されている汎用大型コンピュータシステムの２次記憶装置では、１ドライブ当りの容量が大きいため、他の読み出し／書込み要求に当該ドライブが使用されて、そのドライブを使用できずに待たされることが多く発生した。このタイプのディスクアレイでは汎用大型コンピュータシステムの２次記憶装置で使用されている大容量のドライブを、多数の比較的容量の小さなドライブで構成し、データを分散して格納してあるため、読み出し／書込み要求が増加してもディスクアレイの複数のドライブで分散して処理するため、読み出し／書込み要求がまたされることが減少する。しかし、ディスクアレイは、このように多数のドライブにより構成されるため、部品点数が増加し障害が発生する確率が高くなる。そこで、信頼性の向上を図る目的で、パリティが使用されている。

【０００５】図２１は前記論文において D. Patterson らが提案したＲＡＩＤに述べられている、データを分散して、独立に扱うディスクアレイ（レベル５）内部のデータアドレスを示している。この各アドレスにあるデータは、１回の読み出し／書込みで処理される単位であり、個々のデータは独立している。また、ＲＡＩＤで述べられているアーキテクチャでは、データに対するアドレスは固定されている。前述したようにこのようなシステムでは、信頼性を向上するためパリティを設定することが不可欠である。本システムでは、各ドライブ内の同一アドレスのデータによりパリティが作成される。すなわち、ドライブ＃１から４までのアドレス（１，１）のデータによりパリティが作成され、パリティを格納するドライブの（１，１）に格納される。本システムでは読み出し／書込み処理は、現在の汎用大型計算機システムと同様に、各ドライブに対し当該データをアクセスする。すなわち、図２２に示すように、当該データが格納されているトラックが所属するシリンダの位置とそのシリンダ内において当該データが格納されているトラックを決定するヘッドアドレスＨＨと、そのトラック内のレコードの位置を特定する。具体的には要求データが格納されている当該ドライブ１２の番号（ドライブ番号）と当該ドライブ内のシリンダ番号であるシリンダアドレス（ＣＣ）とシリンダにおいてトラックを選択するヘッド１４０の番号であるヘッドアドレス（ＨＨ）とレコードアドレス（Ｒ）からなるＣＣＨＨＲである。

【０００６】このようなディスクアレイにおいて、例えば図２１のドライブ＃３のアドレス（２，２）のデータを更新する場合、まず、更新される前のドライブ＃３の（２，２）のデータとパリティを格納してあるドライブの（２，２）のパリティを読み出し（ステップ１）、これらと更新する新しいデータとで排他的論理和をとり、新たなパリティを作成する（ステップ２）。パリティの作成完了後、更新する新しいデータをドライブ＃３の（２，２）に、新パリティをパリティを格納するドライブの（２，２）に格納する（ステップ３）。

【０００７】このように、ディスクアレイではデータからパリティを作成しデータと同様にドライブに格納しておく。この時、パリティは、パリティの作成に関与したデータとは別のドライブに格納される。この結果、データを格納したドライブに障害が発生した場合、その障害ドライブ内のデータを復元することが可能となる。

【０００８】これらのディスクアレイでは、現在一般に使用されている汎用大型コンピュータシステムと同様、２次記憶装置内では、個々のデータの格納場所（アドレス）は予め指定したアドレスに固定され、ＣＰＵから当該データへ読み出しまたは書込みする場合、この固定されたアドレスへアクセスすることになっている。

【０００９】このようなレベル５のディスクアレイでは、図２３に示すように、データの格納されているドライブ、パリティの格納されているドライブから古いデータとパリティを読み出すため、ディスクを平均１／２回転待ち、それから読みだしてパリティを作成する。この新しく作成したパリティを書き込むため更に一回転必要となり、データを書き替える場合最低で１．５回転待たなければならない。ドライブにおいては１．５回転ディスクの回転を待つということは非常に大きなオーバヘッドとなる。

【００１０】このような書込み時のオーバヘッドを削減するため、更新後のパリティグループの書込み先のアドレスを動的に変換する方法がＳＴＫ社から出願されている国際出願公開公報ＷＯ ９１／２００７６に開示されている。

【００１１】一方、特開平４－２３０５１２号公報には、ＤＡＳＤアレイのための更新記録方法及び装置が開示されている。この方法によれば、各パリティ・グルー

プのデータ・ブロックとパリティ・ブロックが、障害について関連性のない形式でＤＡＳＤアレイに分散され、かつ各ドライブに複数の論理グループに共通に、未使用スペースが予約される。第１サイクルの間、古いデータ・ブロック及び古いパリティ・ブロックが読み出される。これらのデータと更新用の新しいデータから新しいパリティが計算され、古いデータや古いパリティが保持されていた未使用スペースに、その新しいデータ及び新しいパリティ・ブロックがそれぞれシャドー・ライトされる。

【００１２】

【発明が解決しようとする課題】前記国際出願公開公報ＷＯ ９１／２００７６に開示された動的アドレス変換を行なう場合以下のような問題が生じる。

（１）動的アドレス変換を行なうと、更新前のパリティグループを保持していた領域がそのままドライブ内に無駄な領域として残され、動的にアドレス変換を行なおうとしても、新たな格納領域を確保できなくなる危険性がある。

（２）アドレス情報をテーブル上に管理するとテーブルの容量が増大しメモリのコストが増加し、また、アドレス制御が複雑となるため、処理オーバヘッドが大きくなる。

【００１３】一方、前記特開平４－２３０５１２号公報に記載された未使用スペースを用いる方法によれば、上記（１）の問題はないと考えられる。しかし、上記

（２）のアドレス制御に関しては、その方法が明瞭には開示されていない。さらに、増設時や障害発生時にどのように対処するかについても言及されていない。

【００１４】本発明の目的は、動的アドレス変換を行う場合に、新たな格納領域の確保が容易な、データ書き込み方式を提供することにある。

【００１５】本発明の他の目的は、動的アドレス変換が可能で、かつアドレス制御が簡単であり、増設時や障害回復時にも柔軟に対処できる、データ書き込み方式を提供することにある。

【００１６】

【課題を解決するための手段】本発明によれば、複数のデータと少なくとも１つの誤り訂正コード（例えばパリティコード）を含むパリティグループを構成するドライブ群と、複数の予備の書き込み領域のデータ（ダミーデータ）のドライブ群とにより、１つの論理グループを構成する。この論理グループを構成するドライブ群は、複数のデータとパリティとダミーデータとを各々独立して格納するように構成されている。本発明では、書き込み前と書き込み後では論理グループを構成するドライブ群が異なる点が特徴の１つである。

【００１７】すなわち、書き込み後のデータを書き込むドライブと、書き込み後の新たなパリティを書き込むドライブとを、複数のダミーデータを保有するドライブの中から、書き込み処理全体を見て最も回転待ち時間が少なく、効率的に処理できるように選定する。

【００１８】新しいパリティと新しいデータを格納するドライブを決定した後は、新しいデータを格納するドライブへの書込みが可能になり次第、新しいデータを書込み、旧ドライブ群から書き込み前のデータとパリティを読み出し、書き込む新しいデータとで書き込み後の新しいパリティを作成し、決定済のドライブに書き込み後の新しいデータと新しいパリティを格納する。新しいデータと新しいパリティを格納した後、更新される前のデータが格納されていたドライブと旧パリティが格納されていたドライブとが、各々新しくダミーデータの格納されるドライブとなる。

【００１９】本発明の他の特徴によれば、各論理グループにおいて回転位置情報または回転待ち時間の情報に基づいて書込みデータ（新データ）を格納するドライブを決定する。すなわち、ダミーデータの格納されている複数のドライブ群の中で、どのドライブ内のダミーデータの物理的なアドレスに格納するかを判定する。

【００２０】本発明の他の特徴によれば、論理グループ内のあるドライブに障害が発生した場合、正常な残りのドライブ群内のデータとパリティから、障害が発生したドライブ内のデータまたはパリティを回復処理により復元し、この復元したデータまたはパリティを、予備の書き込み領域に書き込む制御を行う。

【００２１】

【作用】本発明では、１つのパリティグループを構成する複数のデータやパリティ及び複数のダミーデータを、各々独立したドライブに格納する。

【００２２】そのため、書き込みデータ及び書き込み後の新しいパリティを、ダミーデータが格納されていたドライブ群に書き込むことが可能となり、書き込み後のパリティを最小の回転待ち時間で処理できるので、書き込み処理を効率良く処理できる。従来のアレイディスクでは書き込み時に平均１．５回転の回転待ち時間を必要としたのが、平均１回転の回転待ち時間ですむようになる。

【００２３】また、常に複数の予備領域が別々のドライブに確保される構成であるので、ドライブの増設時のアドレス制御や、障害発生時の回復処理も簡単になる。

【００２４】

【実施例】

〔実施例１〕以下本発明の一実施例を図１により説明する。本実施例はＣＰＵ１、アレイディスクコントローラ（以下ＡＤＣ）２，アレイディスクユニット（以下ＡＤＵ）３により構成される。ＡＤＵ３は複数の論理グループ１０により構成され，個々の論理グループ１０はｍ台の磁気記録媒体を用いたＳＣＳＩ (Small Computer System Interface)ドライブ１２と、各々のＳＣＳＩドライブ１２とＡＤＣ２を接続する４つのドライブパス９によ

り構成される。なお、このＳＣＳＩドライブ１２の数は本発明の効果を得るには、特に制限は無い。

【００２５】この論理グループ１０は障害回復単位で、同一の論理グループ内のいずれかのドライブが障害が発生したとき、そのグループ内の他のドライブ内のデータを利用して障害が発生したドライ部内のデータを回復するようになっている。

【００２６】従来のレベル５のディスクアレイでは、各論理グループは、複数のデータを分散して書き込むためのドライブと、それらのデータから生成したパリティを格納するためのドライブからなる。そそれらのデータとパリティとの組はパリティグループと呼ばれる。従って、各論理グループは、分散して書き込むデータの数（これをＮとする）プラス１のドライブからなる。

【００２７】各ドライブは、データ格納用かパリティ格納用ドライブかが定められていた。しかし、本実施例では、各論理グループは少なくともＮ＋３のドライブからなる。したがって、ｍ＝Ｎ＋３である。２つのドライブは、同一のパリティグループ内のデータのその後の更新時に使用する予備の領域（ダミーデータ領域）を提供するのに使用されるのが特徴である。

【００２８】例えば、図３（ａ）に例示するように、ＳＣＳＩドライブ＃１～＃６により１つの論理グループが構成されているとき、データ１～３とパリティＰ１からなる第１のパリティグループは、ドライブ＃１～３、＃５に書き込まれると共に、このパリティグループに対して、予備スペースがドライブ＃４、６に確保される。本実施例では、これらのダミー領域を用いてデータの更新時のドライブの回転待ち時間を減少させるのが特徴である。すなわち、後に詳述するように、このパリティグループ内のデータ例えば、データ１を更新するとき、ことのデータ１とパリティＰ１を読み出し、これらと更新用のデータとから新パリティを生成し、この新パリティと更新用データとをドライブ＃４、＃６という旧データ１と旧パリティＰ１を格納していたドライブと異なるドライブ上のダミー領域に書き込む。

【００２９】旧データ１と旧パリティＰ１を保持していた２つの領域は、新たなこのパリティグループのための新たなダミー領域として使用されることになる。以下、実施例の詳細を説明する。

【００３０】なお、以下の実施例では、パリティコードを用いた例を説明するが、パリティコードに代えて他のハミングコードないしは誤り訂正コードを用いてもよいことは言うまでもない。

【００３１】次にＡＤＣ２の内部構造について図１を用いて説明する。ＡＤＣ２はチャネルパスディレクタ５と２個のクラスタ１３とバッテリバックアップ等により不揮発化された半導体メモリであるキャッシュメモリ７により構成される。チャネルパスディレクタ５は、ＣＰＵ１から発行されたコマンドを取り込む。このキャッシュメモリ７にはドライブ１２に書き込まれ又はそこから読み出されたデータと後述するアドレス変換用テーブル４０（図３）が格納されている。このキャッシュメモリ７およびその中のアドレス変換用テーブルはＡＤＣ２内の全てのクラスタにおいて共有で使用される。クラスタ１３はＡＤＣ２内において独立に動作可能なパスの集合で、各クラスタ１３間においては電源、回路は全く独立となっている。各クラスタ１３にはチャネルパスディレクタ５とキャッシュメモリ７間のパスとして２つのチャネルパス６が設けられ、キャッシュメモリ７と、ＳＣＳＩドライブ１２間のパスとして２つのドライブパス６が、それぞれ２個ずつで構成されている。

【００３２】それぞれのチャネルパス６とドライブパス８はキャッシュメモリ７を介して接続されている。ＣＰＵ１より発行されたコマンドは外部インターフェースパス４を通ってＡＤＣ２のチャネルパスディレクタ５に発行される。本実施例では、ＡＤＣ２は２個のクラスタ１３により構成され、それぞれのクラスタは２個のパスで構成されるため、ＡＤＣ２は合計４個のパスにより構成される。このことから、ＡＤＣ２ではＣＰＵ１からのコマンドを同時に４個まで受け付けることが可能である。

【００３３】図２は図１のチャネルパスディレクタ５と１つのクラスタ１３の内部構造を示した図である。図２に示すように、ＣＰＵ１から外部インターフェースパス４を介してＡＤＣ２に送られてきたコマンドは、チャネルパスディレクタ５内のインターフェースアダプタ（以下ＩＦ　Ａｄｐ）１５の１つにより取り込まれる。各チャネルパス６内のマイクロプロセッサ（以下ＭＰ２０）は外部インターフェースパス４の中でコマンドを取り込んでいるパスがあるかを調べ、そのような外部インターフェースパス４がある場合は、ＭＰ２０はチャネルパススイッチ１６を切り換えてそのパスに取り込まれたコマンドの受付け処理を行なう。もしそのコマンドが受け付けられない場合は受付不可の応答をＣＰＵ１へ送る。１７は信号線、１８はデータパスである。

【００３４】２４はキャッシュアダプタ回路（Ｃ Ａｄｐ）であり、キャッシュメモリ７に対するデータの読み出し、書き込みをＭＰ２０の指示で行う回路で、キャッシュメモリ７の状態の監視、各読み出し、書き込み要求に対し排他制御を行う回路である。Ｃ Ａｄｐ２４により読み出されたデータは、データ制御回路（ＤＣＣ）２２の制御によりチャネルインターフェース回路（ＣＨ ＩＦ）２１に転送される。ＣＨ ＩＦ２１ではＣＰＵ１におけるチャネルインターフェースのプロトコルに変換し、チャネルインターフェースに対応する速度に速度調整する。２７は、データの圧縮回路である。２８は、Ｄｒｉｖｅ ＩＦであり、ＳＣＳＩの読み出し処理手順に従って、読み出しコマンドをドライブユニットパス９を介して発行する。　Ｄｒｉｖｅ ＩＦ２８では転送されてきた当該データをＳＣＳＩドライブ側のキャッシュア

ダプタ回路（Ｃ　Ａｄｐ）１４に転送し、（Ｃ　Ａｄｐ）１４ではキャッシュメモリ７にデータを格納する。３５はデータパスである。３６は、パリティ生成回路であり、キャッシュメモリ７のデータからパリティを生成し、キャッシュメモリに格納する。１２７は、回転位置検出回路である。

【００３５】本実施例では、レベル５のディスクアレイを実現するので、ＣＰＵ１からの複数のＩ／Ｏコマンドで指定されるデータをまとめて、１つのパリティグループに属するデータを作成し、これらをいずれかのＭＰ２０の制御下で、同一の論理グループに書き込む。以下では、説明を簡明にするために、これらのＩ／Ｏコマンドで指定されるドライブ番号は同一とする。この際、これらのデータに対するパリティもＭＰ２０により生成し、そのパリティグループに属するパリティとして同じ論理グループに属する他のドライブに書き込む。

【００３６】図３は、１つの論理グループが６つのつドライブからなる場合について、パリティグループに属するデータがどのように書き込まれるかを示すものである。ここでは、同一のパリティグループに属するデータは３つからなり、これらのデータとこれらのデータから生成されたパリティとが１つのパリティグループを形成する。さらに、このグループに対して２つのダミー領域が確保される。これらはいずれも異なるドライブに属する。すなわち、論理グループは、パリティグループとダミーデータ用に使用されるドライブからなる。

【００３７】このパリティグループのデータ書き込みが最初に行われるときに、ＭＰ２０はいくつかの論理グループ内での同一のドライブ内アドレスを有する空領域をサーチし、そこにこれらの３つのデータとパリテイと２つのダミーデータ用の領域を別々のドライブに確保する。この際、確保された領域のアドレスに関する情報は後に詳述されるアドレス変換テーブル４０に書き込まれる。

【００３８】以上の書き込み動作において従来と異なるのは、あるパリティグループを最初に論理ドライブに書き込むときに、そのパリティグループ用の少なくとも２つの予備領域を確保することにある。

【００３９】本実施例では、ＡＤＵ３を構成するＳＣＳＩドライブ１２は、ＳＣＳＩインターフェースのドライブを使用する。ＣＰＵ１をＩＢＭシステム９０００シリーズのような大型汎用計算機とした場合、ＣＰＵ１からはＩＢＭオペレーティングシステム（ＯＳ）で動作可能なチャネルインターフェースのコマンド体系にのっとってコマンドが発行される。そこで、ＳＣＳＩドライブ１２をＳＣＳＩインターフェースのドライブを使用した場合、ＣＰＵ１からのコマンドを、ＳＣＳＩインターフェースのコマンド体系にのっとったコマンドに変換する必要が生じる。この変換はコマンドのプロトコル変換と、アドレス変換に大きく分けられる。以下にアドレス変換について説明する。

【００４０】ＣＰＵ１から指定されるアドレスは、図２２に示したように、要求データが格納されている当該ドライブ１２の番号（ドライブ番号）と、当該ドライブ内のシリンダ番号であるシリンダアドレス（ＣＣ）と、シリンダにおいてトラックを選択するヘッド１４０の番号であるヘッドアドレス（ＨＨ）と、レコードアドレス（Ｒ）からなる、ＣＣＨＨＲである。

【００４１】従来のＣＫＤフォーマット対応の磁気ディスクサブシステム（例えばＩＢＭ３９９０－３３９０）ではこのアドレスに従ってドライブへアクセスすれば良い。

【００４２】しかし、本実施例では、複数のＳＣＳＩドライブ１２により従来のＣＫＤフォーマット対応の磁気ディスクサブシステムを論理的にエミュレートする。つまり、ＡＤＣ２は複数のＳＣＳＩドライブ１２が、従来のＣＫＤフォーマット対応の磁気ディスクサブシステムで使用されているドライブ１台に相当するようにＣＰＵ１にみせかける。このため、ＣＰＵ１から指定してきたアドレス（ＣＣＨＨＲ）をＳＣＳＩドライブのアドレスにＭＰ２０は変換する。

【００４３】このアドレス変換には図３の（ｂ）に示すようなアドレス変換テーブル４０（以下アドレステーブルとする）が使用される。ＡＤＣ２内のキャッシュメモリ７には、その内部の適当な領域に、アドレステーブル４０が格納されている。本実施例では、ＣＰＵ１が指定してくるドライブは、ＣＫＤフォーマット対応の単体ドライブである。しかし、本発明ではＣＰＵ１は単体と認識しているドライブが、実際は複数のＳＣＳＩドライブにより構成されるため、論理的なドライブ群として定義される。このため、ＡＤＣ２のＭＰ２０はＣＰＵ１より指定してきたアドレス（ＣＰＵ指定ドライブ番号４１とＣＣＨＨＲ４６）を１組の論理グループを構成する。ＳＣＳＩドライブ１２の各々に対するＳＣＳＩドライブ番号とそのＳＣＳＩドライブ内のアドレス（以下ＳＣＳＩ内Ａｄｄｒとする）に変換する。図１に示したように、本実施例では、複数の論理グループを実現できる。アドレステーブル４０は各論理グループに対応して生成される。図３（ｂ）には、図３（ａ）に示したように２つの論理グループがＳＣＡＳＩドライブ＃１～＃６からなると仮定し、その論理グループに対応するアドレステーブルを示す。

【００４４】アドレステーブル４０はＣＰＵ指定ドライブ番号カラム４１とＳＣＳＩドライブアドレスカラム４２とＳＣＡＳＩドライブ情報カラム４３により構成される。これらのカラムの同一行には、同一のパリティグループに属する情報が確保される。ＳＣＳＩドライブアドレスカラム４２はＳＣＳＩドライブ内の実際にデータまたはパリティが格納されているアドレスである、ＳＣＳＩ内Ａｄｄｒを格納するカラム４４と、パリティが格納

されているＳＣＳＩドライブ番号（パリティドライブ番号）を格納するカラム５０と、ダミーデータが格納されているＳＣＳＩドライブ番号（ダミードライブ番号）を格納するカラム５１により構成されている。

【００４５】ＳＣＡＳＩドライブ情報カラム４３は、同一の論理グループに属するＳＣＡＳＩドライブの各々に対応したカラムからなり、それぞれのカラムには、そのドライブに保持されているデータに関連した情報として、ＣＰＵ１から指定された、ドライブ内アドレスＣＣＨＨＲ６と、そのＳＣＡＳＩドライブ内のデータがキャッシュメモリ７内にも存在する場合の、そのデータのキャッシュメモリ７内のアドレス４７と、キャッシュメモリ７がそのデータを保持している場合オン（１）が登録されるキャッシュフラグ４８と、そのドライブにダミーデータが格納されている場合オン（１）となる無効フラグ４９により構成される。なお、図３ではカラム４３内の各カラムには、図３と対比可能なように、それぞれのカラムのそれぞれのエントリがどのデータ及びパリティに関する情報を持つかをそれらのデータまたはパリティで表示してある。

【００４６】このテーブル４０に保持された情報のうち、パリティドライブカラム５０、ダミードライブカラム５１に格納されていないドライブはデータ用のドライブであることがわかる。このテーブル４０の各行のデータは、対応するパリティグループが最初に論理グループに書き込まれたときに、ＭＰ２０により書き込まれる。その後、いずれかのパリティグループ内のデータを更新するＩ／Ｏコマンドが実行されたとき、このテーブルのそのパリテイグループに対する行に属する情報がＭＰ２０により更新される。

【００４７】このテーブル４０のうち、本実施例で特徴的なのはダミードライブ番号が保持されていることである。これは後述する、データの更新時に利用される。このように作成されたテーブル４０は、ＣＰＵ１から論理ドライブ内のデータをアクセスするごとにＭＰ２０により利用される。

【００４８】例えば図３の例において、ＣＰＵ１からＣＰＵ指定ドライブ番号が”Ｄｒｉｖｅ＃１”、ＣＣＨＨＲが”ＡＤＲ８”のデータに対し要求を発行してきた場合、アドレステーブル４０でＣＰＵ指定ドライブ番号がＤｒｉｖｅ＃１に等しい行に属するＳＣＡＳＩドライブ情報カラム４３内エントリを調べ、ＣＣＨＨＲが”ＡＤＲ８”に等しいエントリを探す。図３の例においては、Ｄａｔａ＃２３に対するＣＣＨＨＲが”ＡＤＲ８”となっており、Ｄａｔａ＃２３が要求されたデータであることがわかる。このＤａｔａ＃２３の物理的アドレスはアドレステーブル４０の同じ行からＳＣＳＩドライブ＃２内のＳＣＳＩ内Ａｄｄｒ”ＤＡＤＲ８”であることがわかる。こうして物理的なアドレスへの変換がこのテーブル４２よりなされることがわかる。

【００４９】このアドレス変換の後、ＭＰ２０によりＳＤ＃２のＳＣＳＩドライブ１２のデータ＃２３に対し読み出しまたは書込み要求が発行される。この時アドレステーブル４０においてＤａｔａ＃２３はキャッシュフラグ４８がオフ（０）のため、このデータはキャッシュメモリ７内のＣＡＤＲ２，１に存在しない。もし、キャッシュフラグ４８がオン（１）であればキャッシュメモリ７内には、当該データが存在するので、ドライブ１２からの上記読み出しまたは書き込みは行わず、キャッシュメモリ内のそのデータに対して行われる。また、このデータは無効フラグ４９がオフ（０）のため、このデータはダミーデータでないことがわかる。

【００５０】次に、ＡＤＣ２内での具体的なＩ／Ｏ処理について図１、図２を用いて説明する。ＣＰＵ１より発行されたコマンドはＩＦ　Ａｄｐ１５を介してＡＤＣ２に取り込まれ、ＭＰ１　２０により読み出し要求か書込み要求か解読される。

【００５１】まず、書き込み時は以下のように処理される。但し、以下ではすでにデータが書き込まれた後に行われるデータ更新のためのデータ書き込みについて述べる。

【００５２】ＣＰＵ１からＣＰＵ指定ドライブ番号としてドライブ＃１でＣＣＨＨＲとして”ＡＤＲ８”を指定するデータ更新コマンドが発行されたとする。まず、ＡＤＣ２のＭＰ２０は、ＣＰＵ１からこのコマンドを受け取った後、コマンドを受け取ったＭＰ２０が所属するクラスタ１３内の各チャネルパス６において処理可能かどうかを調べ、可能な場合は処理可能だという応答をＣＰＵ１へ返す。ＣＰＵ１では処理可能だという応答を受け取った後にＡＤＣ２へ書き込みデータを転送する。この転送に先立ち、ＡＤＣ２ではＭＰ２０の指示により、チャネルパスディレクタ５において、チャネルパススイッチ１６がこのコマンドが転送された外部インターフェースパス４とＩＦ　Ａｄｐ１５をチャネルパス６に接続しＣＰＵ１とＡＤＣ２間の接続を確立する。

【００５３】ＣＰＵ１から転送されてきた書き込みデータ（以下新データまたはＮｅｗ　Ｄａｔａとする）はＭＰ２０の指示により、ＣＨ　ＩＦ２１によりプロトコル変換を行ない、外部インターフェースパス４での転送速度からＡＤＣ２内での処理速度に速度調整する。ＣＨ　ＩＦ２１におけるプロトコル変換および速度制御の完了後、データはＤＣＣ２２によるデータ転送制御を受け、Ｃ　Ａｄｐ２４に転送され、Ｃ　Ａｄｐ２４によりキャッシュメモリ７内に格納される。この時、ＣＰＵ１から送られてきた情報が、ＣＰＵ指定アドレスの場合は、読みだしと同様にアドレステーブル７０によりアドレス変換を行い、物理アドレスに変換する。また、ＣＰＵ１から送られてきた情報がデータの場合は、キャッシュメモリ７に格納したアドレスを上記アドレス変換により変換した物理アドレスをキャッシュメモリ７に登録する。この

時、対応するキャッシュフラグ４８をオン（１）とする。

【００５４】この様にキャッシュメモリ７に新データを格納したのをＭＰ２０が確認したら、ＭＰ２０は書込み処理の完了報告をＣＰＵ１に対し報告する。

【００５５】なお、キャッシュメモリ７内に保持されている新データに対し、さらに書き込み要求がＣＰＵ１から発行された場合は、キャッシュメモリ７内に保持されている新データを書き替える。

【００５６】キャッシュメモリ７に新データが格納された後は、以下のようにして新しくパリティを更新し、論理グループ１０内のＳＣＳＩドライブ１２へ新データと新パリティが格納される。

【００５７】図２に戻って、ＭＰ２０はアドレステーブル４０を参照したときに、データ、ダミーデータ、パリティが格納されているＳＣＳＩドライブ番号を認識する。

【００５８】ＭＰ２０は、Ｄｒｉｖｅ ＩＦ２８に対し、該当するドライブ１２に対する書き込み処理を行なうように指示する。

【００５９】ここでの書き込み処理は、論理グループ１０において新データを実際に書き込む処理と、新データの書き込みによりパリティを新たに作り直す（以下新パリティ、Ｎｅｗ Ｐａｒｉｔｙとする）ための処理すなわち、このパリティを作成するための書き込み前のデータ（以下旧データ、Ｏｌｄ Ｄａｔａとする）と書き込み前のパリティ（以下旧パリティとする）を読み出し、新パリティを作成する処理と、新データと新パリティを書き込む一連の処理からなる。

【００６０】図４に示すようにＣＰＵ１からＳＤ＃１のＳＣＳＩドライブ１２のデータ＃１に書き込み要求が発行された場合、ＭＰ２０は、ＳＤ＃１ドライブが所属する論理グループ１０において、ドライブＳＤ＃１とダミーデータが格納されているドライブＳＤ＃４，６とパリティが格納されているドライブＳＤ＃５に対し、使用権の獲得を行なう。

【００６１】その後は、図５のフローチャートに示すような処理を行なっていく。まず、ＭＰ２０は回転位置検出回路（ＲＰＣ）１２７によりドライブＳＤ＃１、４、５、６における回転位置を検出する（ステップ５０２）。ＲＰＣ１２７により回転位置を検出した後にＭＰ２０は各ＳＣＳＩドライブの回転待ち時間を算出する（５０４）。

【００６２】図６に示すように、論理グループ１０を構成する各ドライブＳＤ＃１～ＳＤ＃６には旧データ（Ｄａｔａ＃１～＃３）及びダミーデータ（Ｄｕｍｍｙ）、旧パリティ（Ｐａｒｉｔｙ＃１）が格納されている。この論理グループ１０内のデータ１００は、ディスク１３０上では図６の下段に示すようになっており、各ドライブＳＤ＃１～ＳＤ＃６によって当該データ１００がヘッド１４０の下に来るまでのディスク１３０の回転を待つ時間が異なっている。このディスク１３０の回転を待つ時間のことを回転待ち時間（ＴＷ）という。

【００６３】回転待ち時間ＴＷを算出した後に、以下のように新パリティと新データを格納するＳＣＳＩドライブを決定する。図５のフローチャートに示すように、まず、ＭＰ２０は旧データ（Ｏｌｄ Ｄａｔａ）が格納されているＳＣＳＩドライブと、旧パリティ（Ｏｌｄ Ｐａｒｉｔｙ）が格納されているＳＣＳＩドライブのうち、回転待ち時間ＴＷの大きい方を判定する（５０６～５１２）。

【００６４】図７の（ａ）の例では、旧データ（Ｏｌｄ Ｄａｔａ）が格納されているドライブＳＤ＃１の回転待ち時間がＴＷ１、旧パリティ（Ｏｌｄ Ｐａｒｉｔｙ）が格納されているドライブＳＤ＃５の回転待ち時間ＴＷ５より大きい。そこで、ＭＰ２０はドライブＳＤ＃１のを基準に新データ（Ｎｅｗ Ｄａｔａ）と新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）を書き込むドライブを決定する。具体的にはＭＰ２０はＲＰＣ１２７からの情報によりダミーデータが格納されているドライブＳＤ＃４，６において、ドライブＳＤ＃１から書き込み前の旧データ（Ｏｌｄ Ｄａｔａ）を読み出し（５０６）、新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）の作成後に早く書き込むことが可能な方のドライブを判定し、このドライブに新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）を格納することに決定する（５１２）。一方新データ（Ｎｅｗ Ｄａｔａ）は残りのダミーデータが格納されているドライブに格納することに決定する（５１４）。このようにして、新データ（Ｎｅｗ Ｄａｔａ）と新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）を格納するＳＣＳＩドライブ１２を決定する。

【００６５】図７（ａ）の例では、ドライブＳＤ＃１から書き込み前の旧データを読み出した後は、ドライブＳＤ＃６の方がドライブＳＤ＃４よりも早く書き込める。（ＳＤ＃４のＳＣＳＩドライブに書き込むためには、１回転待たなければならない。）そのため、ドライブＳＤ＃６を書き込み後の新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）を格納するドライブとし、ドライブＳＤ＃４は、書き込む新データ（ＮｅｗＤａｔａ）を格納するドライブとする。

【００６６】また、図７（ｂ）の例では、ドライブＳＤ＃１から書き込み前の旧データ（Ｏｌｄ Ｄａｔａ）を読み出した後、ドライブＳＤ＃４の方がドライブＳＤ＃６よりも早く書き込めるため、ドライブＳＤ＃４を書き込み後の新パリティ（ＮｅｗＰａｒｉｔｙ）を格納するドライブとし、ドライブＳＤ＃６は書き込む新データ（Ｎｅｗ Ｄａｔａ）を格納するドライブとする。

【００６７】書き込み後の新データ（Ｎｅｗ Ｄａｔａ）および新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）を格納するドライブを決定した後は、図５のフローチャートに従って以下のように処理する。以下の書き込み処理の説

明は、図７の（ａ）を例に行なう。

【００６８】ＭＰ２０はＤｒｉｖｅ ＩＦ２８に新データ（Ｎｅｗ Ｄａｔａ）をドライブＳＤ＃４に書き込むよう指示する。Ｄｒｉｖｅ ＩＦ２８ではＳＣＳＩの書込み手順にしたがってドライブユニットパス９の中の１本を介してドライブＳＤ＃４に書込みコマンドを発行する。

【００６９】同時にＭＰ２０はドライブＳＤ＃１に対して旧データ（Ｏｌｄ Ｄａｔａ）の読み出しと、ドライブＳＤ＃５に対して旧パリティ（Ｏｌｄ Ｐａｒｉｔｙ）の読み出しとドライブＳＤ＃６への新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）の書込み要求を発行するようにＤｒｉｖｅ ＩＦ２８に指示する。

【００７０】Ｄｒｉｖｅ ＩＦ２８から読み出しまたは書込みコマンドを発行されたこれらのドライブでは、Ｄｒｉｖｅ ＩＦ２８から送られてきたＳＣＳＩ内Ａｄｄｒ４４へシーク、回転待ちのアクセス処理を行なう。ドライブＳＤ＃４においてアクセス処理が完了し書込みが可能になり次第、Ｃ Ａｄｐ１４はキャッシュメモリ７から書き込む新データ（Ｎｅｗ Ｄａｔａ）を読み出してＤｒｉｖｅ ＩＦ２８へ転送し、Ｄｒｉｖｅ ＩＦ２８では転送されてきた新データ（Ｎｅｗ Ｄａｔａ）をドライブユニットパス９の中の１本を介してＳＤ＃４のドライブへ転送する。新データ（Ｎｅｗ Ｄａｔａ）のドライブＳＤ＃４への書込みが完了すると（５１０）、ドライブＳＤ＃４はＤｒｉｖｅ ＩＦ２８に完了報告を行ない、Ｄｒｉｖｅ ＩＦ２８がこの完了報告を受け取ったことを、ＭＰ２０に報告する（５１６）。

【００７１】ＳＤ＃４のドライブに対する新データ（Ｎｅｗ Ｄａｔａ）の格納は完了しても、キャッシュメモリ７内には新データ（Ｎｅｗ Ｄａｔａ）が存在しており、パリティの更新はキャッシュメモリ７内に格納されている新データ（Ｎｅｗ Ｄａｔａ）で行なう。

【００７２】ドライブＳＤ＃１，５においてもアクセス処理が完了し旧データ（Ｏｌｄ Ｄａｔａ）および旧パリティ（Ｏｌｄ Ｐａｒｉｔｙ）の読み出しが可能になり次第、旧データ（Ｏｌｄ Ｄａｔａ）および旧パリティ（Ｏｌｄ Ｐａｒｉｔｙ）を読み出し、キャッシュメモリ７に格納する（５１２）。ドライブＳＤ＃１から旧データ（Ｏｌｄ Ｄａｔａ）を、ドライブＳＤ＃５から旧パリティ（Ｏｌｄ Ｐａｒｉｔｙ）を読み出し、それぞれをキャッシュメモリ７に格納した後、ＭＰ２０はキャッシュメモリ７内に格納されている書き込む新データ（Ｎｅｗ Ｄａｔａ）とで排他的論理和により、更新後の新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）を作成するようにパリテイ生成回路（ＰＧ）３６に指示を出し、ＰＧ３６において新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）を作成しキャッシュメモリ７に格納する（５１８）。

【００７３】新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）をキャッシュメモリ７に格納した後、ＭＰ１ ２０では新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）の作成が完了したことを認識し、ドライブＳＤ＃６に、更新後の新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）を書き込むようにＤｒｉｖｅ ＩＦ２８に対し指示する。ドライブＳＤ＃６への更新後の新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）の書込み方法は、先に述べた書き込む新データ（Ｎｅｗ Ｄａｔａ）をドライブＳＤ＃４に書き込んだ方法と同じである。

【００７４】ドライブＳＤ＃６では既に、Ｄｒｉｖｅ ＩＦ２８から書込みコマンドが発行されており、指示ＳＣＳＩ内Ａｄｄｒ４４へシーク、回転待ちのアクセス処理を行なっている（５２０）。新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）は既に作成されキャッシュメモリ７に格納されており、ドライブＳＤ＃６におけるアクセス処理が完了した場合、Ｃ Ａｄｐ１４はキャッシュメモリ７から新パリティ（ＮｅｗＰａｒｉｔｙ）を読み出してＤｒｉｖｅ ＩＦ２８へ転送する。Ｄｒｉｖｅ ＩＦ２８では転送されてきた新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）をドライブユニットパス９の中の１本を介してドライブＳＤ＃６へ転送する（５２２）。

【００７５】新データ（Ｎｅｗ Ｄａｔａ）及び新パリティ（Ｎｅｗ Ｐａｒｉｔｙ）の当該ドライブＳＤ＃４，６への書込みが完了すると、当該ドライブはＤｒｉｖｅ ＩＦ２８に完了報告を行ない、Ｄｒｉｖｅ ＩＦ２８がこの完了報告を受け取ったことを、ＭＰ２０に報告する。

【００７６】この時、ＭＰ２０は、この新データ（Ｎｅｗ Ｄａｔａ）をキャッシュメモリ７上に残さない場合は、この報告を元にアドレステーブル４０のキャッシュフラグ４８をオフ（０）にする。さらに、アドレステーブル４０において、書き込み前の旧データ（Ｏｌｄ Ｄａｔａ）の論理アドレス４５の無効フラグをオン（１）とし、書き込み前の論理アドレス４５内のＣＣＨＨＲ４６のアドレスを、書き込み後の論理アドレス４５のＣＣＨＨＲ４６に登録し、無効フラグをオフ（０）とする。

【００７７】また、キャッシュメモリ７内に書き込む新データ（Ｎｅｗ Ｄａｔａ）を保持する場合は、書き込み後のデータに対応するエントリ内のキャッシュアドレス４７に、キャッシュメモリ７内の新データ（Ｎｅｗ Ｄａｔａ）が格納されているアドレスを登録し、キャッシュフラグ４８をオン（１）とする。さらに、書き込みを行った論理グループ１０内のＳＣＳＩ内Ａｄｄｒ４４に対し、パリティドライブ番号５０とダミードライブ番号５１を書き込み後のＳＣＳＩドライブ番号４３に変更する。

【００７８】図７の（ｂ）の例についても同様に処理することが可能である。

【００７９】以上述べたようにして書込み処理を行なうが、その時の回転位置検出方法について以下に詳しく説明する。

【００８０】回転待ち時間の算出方法は、図８に示すよ

うにディスク１３０の面上をインデックスを基準として扇型に等分割する。本実施例では図８に示すように１６分割した。なお、この分割数に制限が無いことは明らかである。この等分割した各領域には番号（１～１６）が付けられており、各領域の開始地点にその番号が記録されている。回転待ち時間ＴＷの算出は、各ＳＣＳＩドライブ１２において現在ヘッド１４０の下にある領域の番号をＲＰＣ１２７に送り、ＲＰＣ２７では当該データの格納されている領域１００（アドレスでも可）との差から回転待ち時間ＴＷを計算する。図８を例に計算方法を示す。ヘッド１４０は領域１３の開始地点に位置している。要求データは領域２、３に存在している。要求データの開始領域は領域２である。そこで、ディスク１３０が一回転する時間を１６．６（ｍｓ）とすると回転待ち時間ＴＷは以下のように求められる。

【００８１】（１）（ヘッドの位置する領域番号－要求データの開始領域番号）＞０なら
回転待ち時間＝（総分割数－（ヘッドの位置する領域番号－要求データの開始領域番号））／総分割数×一回転時間＝（１６－（１３－２））／１６×１６．６（ｍｓ）
また、要求データの開始領域を領域１５とした場合は、
（２）（ヘッドの位置する領域番号－要求データの開始領域番号）≦０、
回転待ち時間＝（要求データの開始領域番号－ヘッド位置領域番号）／総分割数×一回転時間＝（１５－１３）／１６×１６．６（ｍｓ）
となる。

【００８２】なお、回転待ち時間を時間ではなく、ディスク面上におけるヘッドの位置する領域である回転位置情報そのものから要求データの開始領域までの領域数で考えても問題は無い。

【００８３】図９は、図７の（ａ）の場合での書き込み前の旧データ（Ｏｌｄ　Ｄａｔａ）と旧パリティ（Ｏｌｄ　Ｐａｒｉｔｙ）、書き込み後の新データと新パリティのデータの流れを示している。書込み処理を行なった結果、書き込み前に図１０の（Ａ）に示すように、旧データ（Ｏｌｄ　Ｄａｔａ）が格納されていたＳＤ＃１のＳＣＳＩドライブ１２と旧パリティ（Ｏｌｄ　Ｐａｒｉｔｙ）が格納されていたＳＤ＃５のＳＣＳＩドライブ１２は、図１０の（Ｂ）に示すように、書き込み後はダミーデータが格納されているＳＣＳＩドライブ１２となる。この時、ＳＤ＃１のＳＣＳＩドライブ１２には書き込み前の旧データ（Ｏｌｄ　Ｄａｔａ）が格納されており、ＳＤ＃５のＳＣＳＩドライブ１２には書き込み前の旧パリティ（Ｏｌｄ　Ｐａｒｉｔｙ）が格納されているが、これらは、意味の無いデータまたはパリティでダミーデータとし、書き込み前と書き込み後では論理グループ１０においてパリティを作成するパリティグループを構成するＳＣＳＩドライブ１２が異なる。

【００８４】書き込み処理においては、新パリティ（Ｎｅｗ　Ｐａｒｉｔｙ）を書き込む処理により書き込み処理全体の処理時間が左右される。このように、書き込み後の新データ（Ｎｅｗ　Ｄａｔａ）及び新パリティ（Ｎｅｗ　Ｐａｒｉｔｙ）を、書き込み前の旧データ（Ｏｌｄ　Ｄａｔａ）及び旧パリティ（Ｏｌｄ　Ｐａｒｉｔｙ）が格納されているＳＣＳＩドライブ１２ではなくダミーデータが格納されているＳＣＳＩドライブ１２に書き込むことにより、データの書き込みにより必要となるパリティの変更を行なった後に、最も早く書き込めるＳＣＳＩドライブ１２に書き込むことが可能となり、従来のアレイディスクでは図２３に示すように書き込み時に平均１．５回転の回転待ち時間が必要としたのが、平均１回転の回転待ち時間で済むため書き込み処理全体の処理時間を短縮できる。

【００８５】次に読み出し要求の場合の処理方法を以下に示す。

【００８６】ＭＰ２０が読み出し要求のコマンドを認識すると、ＭＰ２０はＣＰＵ１から送られてきたＣＰＵ指定ドライブ番号とＣＣＨＨＲ（以下両方を併せてＣＰＵ指定アドレスとする）によりアドレステーブル４０を参照し、要求されたデータに対する物理アドレスへの変換を行ない、それとともにそのデータがキャッシュメモリ７内に存在するかどうかをキャッシュフラグ４８で判定する。キャッシュフラグ４８がオンでキャッシュメモリ７内に格納されている場合（キャッシュヒット）は、ＭＰ２０がキャッシュメモリ７から当該データを読みだす制御を開始し，キャッシュメモリ７内に無い場合（キャッシュミス）は当該ドライブ１２へその内部の当該データを読みだす制御を開始する。

【００８７】キャッシュヒット時、ＭＰ２０はアドレステーブル４０によりＣＰＵ１から指定してきたＣＰＵ指定アドレスをキャッシュメモリ７のアドレスに変換しキャッシュメモリ７へ当該データを読み出しに行く。具体的にはＭＰ２０の指示の下で、キャッシュアダプタ回路（Ｃ　Ａｄｐ）２４によりキャッシュメモリ７から当該データが読み出される。Ｃ　Ａｄｐ２４はキャッシュメモリ７に対するデータの読み出し、書き込みをＭＰ２０の指示で行う回路で、キャッシュメモリ７の状態の監視、各読み出し、書き込み要求に対し排他制御を行う回路である。

【００８８】Ｃ　Ａｄｐ２４により読み出されたデータは、データ制御回路（ＤＣＣ）２２の制御によりチャネルインターフェース回路（ＣＨ　ＩＦ）２１に転送される。ＣＨ　ＩＦ２１ではＣＰＵ１におけるチャネルインターフェースのプロトコルに変換し、チャネルインターフェースに対応する速度に速度調整する。ＣＨ　ＩＦ２１におけるプロトコル変換および速度調整の後は、チャネルパスディレクタ５において、チャネルパススイッチ１６が外部インターフェースパス４を選択し、ＩＦ　Ａ

ｄｐ１５によりＣＰＵ１へデータ転送を行なう。

【００８９】一方、キャッシュミス時、ＭＰ２０はアドレス変換テーブル４０を用いて、読み出すべきデータが属するドライブとその中のアドレスを判別し、ＤｒｉｖｅＩＦ２８に対し、当該ドライブ１２への読み出し要求を発行するように指示する。Ｄｒｉｖｅ　ＩＦ２８ではＳＣＳＩの読み出し処理手順に従って、読み出しコマンドをドライブユニットパス９を介して発行する。Ｄｒｉｖｅ　ＩＦ２８から読み出しコマンドを発行された当該ＳＣＳＩドライブ１２においては、指示されたＳＣＳＩ内Ａｄｄｒ４４へシーク、回転待ちのアクセス処理を行なう。当該ＳＣＳＩドライブ１２におけるアクセス処理が完了した後、当該ＳＣＳＩドライブ１２は当該データを読み出しドライブユニットパス９を介してＤｒｉｖｅ　ＩＦ２８へ転送する。

【００９０】Ｄｒｉｖｅ　ＩＦ２８では転送されてきた当該データをＳＣＳＩドライブ側のキャッシュアダプタ回路（Ｃ　Ａｄｐ）１４に転送し、（Ｃ　Ａｄｐ）１４ではキャッシュメモリ７にデータを格納する。この時、Ｃ　Ａｄｐ１４はキャッシュメモリ７にデータを格納することをＭＰ２０に報告する。ＭＰ２０はこの報告を元にアドレステーブル４０内の、ＣＰＵが読み出し要求を発行したＣＰＵ指定アドレスに対応したエントリ４５のキャッシュフラグ４８をオン（１）にし、さらにキャッシュアドレス４７にキャッシュメモリ７内のデータを格納したアドレスを登録する。キャッシュメモリ７にデータを格納し、アドレステーブル４０のこのデータに対応するエントリ４５のキャッシュフラグ４８をオンにし、キャッシュメモリ７内アドレス４７を更新した後は、キャッシュヒット時と同様な手順でＣＰＵ１へ当該データを転送する。

【００９１】本実施例では、パリティは論理グループ１０を構成する各ＳＣＳＩドライブ１２において、各ＳＣＳＩ内Ａｄｄｒが同一のデータに対して作成され、そのパリティもデータと同一のＳＣＳＩ内Ａｄｄｒ４４に格納されるとして説明してきた。この方法によれば、ドライブ番号のみでアドレス管理できる利点がある。

【００９２】しかし、図１０に示すように、パリティ・データが同一ＳＣＳＩ内Ａｄｄｒ４４ではなく、アドレスが１ずつずれていても同様に実現できる。この場合、各データについて、ドライブ番号のみならずドライブ内アドレスの管理も必要となってくる。システム全体の管理の難易度を考慮して、このずらす量を選定すればよい。

【００９３】〔実施例２〕本発明の第２の実施例を図１１、図１２を用いて説明する。本実施例では、実施例１で示したシステムにおいてＳＣＳＩドライブ１２に障害が発生した時に、その障害が発生したＳＣＳＩドライブ１２内のデータを回復し、それを格納するための領域にダミーデータの領域を使用する例を示す。論理グループ１０内の各ＳＣＳＩドライブ１２では、その内部の各々対応する同一ＳＣＳＩ内Ａｄｄｒ４４のデータによりパリティグループを構成する。具体的には図１１に示すように、ドライブＳＤ＃１，２，３内のデータ＃１、２、３でパリティ生成回路ＰＧ３６によりパリティ＃１が作られ、ドライブＳＤ＃５内に格納される。本実施例ではパリティは奇数パリティとし、各々対応するビットについて１の数を数え、奇数であれば０、偶数であれば１とする（排他的論理和）。

【００９４】もし、ドライブＳＤ＃１に障害が発生したとする。この時、データ＃１は読み出せなくなる。

【００９５】そこで、図１２に示すように、残りのデータ＃２、３とパリティ＃１をキャッシュメモリ７に転送し、ＭＰ２０はＰＧ３６に対しデータ＃１を復元する回復処理を早急に行なうように指示する（１２０２）。この回復処理を行ないデータ＃１を復元した後（１２０６～１２０８）、ＭＰ１　２０は、アドレステーブル４０により、論理グループ内の同一ＳＣＳＩ内Ａｄｄｒ４４（ＤＡＤＲ１）におけるダミーデータの格納されてダミードライブ番号５１（ＳＤ＃４，６）を認識し、このデータ＃１をＳＤ＃４または６のどちらかのダミーデータのアドレスに格納する（１２１０～１２１２）。

【００９６】本実施例では、ダミーデータの領域に回復処理により復元されたデータ＃１を格納する。これにより、ダミーデータ領域を実施例１で示したような、書き込み処理時の回転待ち時間を短縮させるためだけではなく、ＳＣＳＩドライブ１２に障害が発生したときに、復元したデータを格納するためのスペアー領域としても活用する。

【００９７】この様に、ＭＰ１　２０がダミーデータ領域に回復したデータ＃１を格納した後は、キャッシュメモリ７にある図３に示すアドレステーブル４０において、ダミードライブ番号５１の中で、回復データを格納した方のダミードライブ番号を削除し、この削除したドライブ番号に対するエントリ４５に、回復したデータ＃１の元のエントリ４５の内容を複写する（１２１４）。

【００９８】図１１に示すように、ドライブＳＤ＃１には、データ＃１の他にダミーデータ、パリティ、データ＃１３、１６、１９、２２が格納されている。ダミーデータについては回復処理を行ない復元する必要は無い。パリティ＃３はドライブＳＤ＃３，４，５からデータ＃７、８、９を読み出して新たに作成しドライブＳＤ＃２か６のダミーデータ領域に格納する。データ＃１３はドライブＳＤ＃３，５，６からパリティ、データ＃１４、１５を読み出して、回復処理を行ない復元し、ドライブＳＤ＃２または４のダミーデータ領域に格納する。以下同様に他のデータについても回復処理を行ない論理グループ１０内のダミーデータ領域に格納していく。

【００９９】なお、ドライブＳＤ＃２，３，４，５，６のダミーデータ領域に、ドライブＳＤ＃１の回復データ

を全て格納した後は、ダミーデータが論理グループ１０において一つしかないため、実施例１で述べたような書き込み時の回転待ちを短くすることは出来ないため、従来のアレイディスクであるＲＡＩＤのレベル５の処理となる。また、ドライブＳＤ＃１の回復データを全て格納した後は、ドライブＳＤ＃２，３，４，５，６の中で更にもう一台のドライブに障害が発生した場合、同様にその障害が発生したドライブ内のデータについて回復処理を行ない、残りのダミーデータ領域に格納し、処理を行なえる。

【０１００】この様にして、論理グループ１０内のダミーデータ領域を全て使いきってしまった場合は、障害ＳＣＳＩドライブ１２を正常のＳＣＳＩドライブ１２に交換し、この交換した正常なＳＣＳＩドライブ１２は全てダミーデータ領域として論理グループを再構成する。障害ＳＣＳＩドライブ１２を正常のＳＣＳＩドライブ１２に交換した直後は、ダミーデータ領域が特定ＳＣＳＩドライブ１２に集中した形になっているため、このＳＣＳＩドライブ１２が使用出来ずにまたされることが多くなりネックとなるため、実施例１で示した回転待ち時間を短縮する効果が、効率的に発揮出来ない。

【０１０１】しかし、時間が立つにつれて、ダミーデータが分散されて、ＳＣＳＩドライブ障害前の状態に戻っていき、次第に解消されていく。もし、この時間が問題となる場合は、ＳＣＳＩドライブに障害が発生したことを感知した場合、正常なＳＣＳＩドライブに交換して、この交換した正常なＳＣＳＩドライブに障害が発生したＳＣＳＩドライブ内のデータとパリティをユーザが復元することも可能とする。なお、この時ダミーデータに関しては復元せずにダミーデータ領域として空けておく。

【０１０２】本実施例ではこの回復処理と、ダミーデータ領域へ復元したデータを書き込む処理をＭＰ２０が自動的に行なう。この様に自動的に行なうことによりＳＣＳＩドライブに障害が発生した場合、障害が発生したＳＣＳＩドライブを正常なＳＣＳＩドライブに交換し回復したデータを書き込むのと比較し、本発明ではシステムを使用するユーザがＳＣＳＩドライブに障害が発生するとすぐに正常なＳＣＳＩドライブと交換する必要が無いため、ユーザの負担が軽くなる。

【０１０３】〔実施例３〕本発明の第３の実施例を図１３～図１５により説明する。本実施例では、実施例１、２において示したディスクアレイシステムの論理グループ１０において、要求性能に合わせてダミーデータの数を増やすことを可能とする。ここでいう要求性能とは読み出し／書込み処理を高速に行ない処理時間を短縮させるアクセス性能と、障害発生時に対処する信頼性能の両方をいう。ダミーデータの数を増加させる方法は、図１３のように全てがダミーデータにより構成されている拡張ドライブを増設する。

【０１０４】一例として図１４に示すように、論理グループ１０内においてダミーデータの数を３個にした場合を示す。この例では、ＳＤ＃１、２、３のＳＣＳＩドライブ１２にデータが格納されており、ＳＤ＃４，６及びＳＤ＃７（拡張ドライブ）のＳＣＳＩドライブ１２にはダミーデータが格納されており、ＳＤ＃５のＳＣＳＩドライブ１２にはパリティが格納されている。

【０１０５】本実施例でのデータの更新時の処理は、実施例１に比べて、ダミードライブの数が多い点が異なる。つまり新パリティを書き込むドライブを選択するとき、３つのダミードライブから実施例１の方法で選択する。新データの書き込みは残りの２つのダミードライブのうち、より早く書き込めるドライブを選択する。図１４はそのときの動作の一例である。

【０１０６】本実施例のようにダミーデータを増加させることにより、書き込み処理における選択対象のドライブが増え、書込み時のオーバヘッド（回転待ち）を減少させることが可能となる。また、障害が発生した場合、実施例２のようにその障害が発生したデータを回復して、この回復したデータを格納することが可能であり、この回復したデータ格納する領域を増加させることにより信頼性の向上を図ることが可能となる。高性能な処理を要求するデータはこの様にダミーデータを多く持つようにする。また、一つの論理グループ１０内においてダミーデータの数が違うグループ混在させることも可能である。

【０１０７】ＭＰ２０は当該論理グループ１０内において、どのデータ、ダミーデータ、パリティがどの様に構成されているか、キャッシュメモリ７内のアドレステーブル４０により認識することが可能である。このため、ＣＰＵ１から読み出しまたは書込み要求が発行されてきた場合、当該データが格納されている論理グループ１０の構成を、アドレステーブル４０を調べることで認識し、実施例１のように処理するか、実施例３のように処理するかを、ＭＰ２０が判断し、処理を行なう。

【０１０８】なお、本実施例では実施例２の効果を実現することが可能であることは明らかである。また、本実施例ではダミーデータ数を３としたが、この数は本発明を制約しない。

【０１０９】本実施例では、図１３に示すように、拡張ドライブにはダミーデータのみが格納されているとして、論理グループ１０における同一ＳＣＳＩ内Ａｄｄｒ４４でダミーデータ数を増加した。

【０１１０】このように、ダミーデータ数を増加させた後、容量または、性能の向上を図るため、ダミーデータの一部に、新規にデータを格納する方法を図１５で説明する。図１５は、ダミーデータの追加処理（Ａ）と、容量増設に伴う新規データの書き込み処理（Ｂ）とからなっている。ダミーデータの追加処理（Ａ）では、図１３に示したように、６台のＳＣＳＩドライブ１２により構成されていた論理グループ内に、１台のＳＣＳＩドライ

ブ１２を増設し、これを拡張ドライブとした場合、先に述べたように、この拡張ドライブは、ダミーデータにより構成されているとし、論理グループ１０における同一ＳＣＳＩ内Ａｄｄｒ４４でダミーデータ数を増加させる。この時、アドレステーブル４０のＳＣＳＩドライブ番号４３に拡張ドライブのＳＣＳＩドライブ番号を追加し、ダミードライブ番号５１に拡張ドライブのＳＣＳＩドライブ番号を追加する。また、追加したＳＣＳＩドライブ番号の各ＳＣＳＩ内Ａｄｄｒ４４に対応する全論理アドレス４５において無効フラグ４９をオン（１）として全てをダミーデータとする（１５０２）。

【０１１１】次に新規データの書き込み処理（Ｂ）を説明する。もし、ユーザから容量の増設要求がＡＤＣ２のＭＰ１　２０に発行され、新規データを書き込む場合、ＭＰ１　２０はアドレステーブル４０により論理グループ１０における同一ＳＣＳＩ内Ａｄｄｒ４４の３個のダミーデータのアドレスを認識し、現在、他の読み出し／書き込み処理を行っていないＳＣＳＩドライブ１２に格納されてダミーデータを１個選択する（１５０４）。ＭＰ１　２０によるダミーデータの選択が完了したら、ＭＰ１　２０は、新規データをこの選択したダミーデータのアドレスに通常の書き込み処理と同様に書き込む（１５０６）。なお、この時、パリティの更新は、旧データはダミーデータが全て０として、旧データの読み出しは不要である。また、新パリティの書き込みについては、新規データを書き込んだダミーデータ以外の残りの２個のダミーデータの中で、どちらか早く書き込める方に書き込んでもかまわない。

【０１１２】以上のように、ダミーデータに新規データを書き込んだ後、ＭＰ１　２０はアドレステーブル４０において、新規データが書き込まれたダミーデータの論理アドレス４５を新規データに変更し、ダミードライブ番号カラム５１から新規データが書き込まれたダミーデータのＳＣＳ１ドライブ番号を削除する（１５０８）。以下同様に新規データを次々に書き込み、容量の増設を図る。

【０１１３】このように、論理グループ１０に分散されているダミーデータに新規データを分散して書き込むため、拡張したドライブに対する新規データの書き込みが、集中しないため、効率良く処理することが可能となる。

【０１１４】〔実施例４〕本発明の第４の実施例を図１６で説明する。本実施例では複数の論理グループ１０間に渡ってダミーデータを共有することを可能にする。ディスクアレイサブシステムは、図１６に示すように複数の論理グループ１０により構成される。本実施例では実施例３で述べたような、高性能を要求するデータについてダミーデータを増加させるのを、論理グループ１０内のＳＣＳＩドライブ１２の数を増加させるのではなく、論理グループ１０間のＳＣＳＩドライブ１２の割り当て

方を可変とすることで実現する。本実施例では、サブ論理グループはパリティグループと、このパリティグループに対応するダミーデータにより構成されるグループと定義し、論理グループ１０はこのサブ論理グループにより構成されるとする。

【０１１５】図１６においてサブ論理グループ＃１とサブ論理グループ＃２においてＳＤ＃７のＳＣＳＩドライブ１２のダミーデータを共有する場合を以下に示す。サブ論理グループ＃１においてはドライブＳＤ＃１のデータ＃１、ドライブＳＤ＃２のデータ＃２、ＳＤ＃３のデータ＃３とＳＤ＃４，６のダミーデータとＳＤ＃５のパリティにより構成されている。サブ論理グループ＃１についてはドライブＳＤ＃７のダミーデータをサブ論理グループ＃２と共有し拡大論理グループ＃１とする。

【０１１６】拡大論理グループ＃１ではダミーデータが３個あり、サブ論理グループ＃２ではダミーデータは２個である。一方、サブ論理グループ＃２においてはドライブＳＤ＃８のデータ＃４、ＳＤ＃１０のデータ＃５、ＳＤ＃１１のデータ＃６、ＳＤ＃７，９のダミーデータと、ＳＤ＃１２のパリティ＃２により構成されている。

【０１１７】拡大論理グループ＃１では実施例２、３での制御が行なわれ、サブ論理グループ＃２では実施例１、２の制御が行なわれる。ＭＰ２０は当該サブ論理グループが、どのＳＣＳＩドライブ１２内のどのデータとパリティとダミーデータにより構成されているか、キャッシュメモリ７内のアドレステーブル４０により認識することが可能である。アドレステーブル４０にはＣＰＵ１が指定してきたＣＰＵ指定アドレスが、どのサブ論理グループに属するものなのかのテーブルを持っている。このため、ＣＰＵ１から読み出しまたは書込み要求が発行されてきた場合、アドレステーブル４０により当該データが、どのサブ論理グループに所属するか、ＭＰ２０が認識することが可能である。

【０１１８】以上のように本実施例ではサブ論理グループ間においてダミーデータを共有し拡大論理グループを構成することが可能となる。拡大論理グループにおいて、どのサブ論理グループがダミーデータを共有するかはユーザは自由に指定可能である。この時、サブ論理グループ間において、ＣＰＵ１からの読み出しまたは書込み要求が競合しないような論理グループ１０でダミーデータを共有するようにする方が良い。

【０１１９】なお、各パリティグループごとに予備領域を持つことは、管理を単純化する効果があるが、用途によっては、パリティグループごとに対応させる必要はない。

【０１２０】すなわち、各パリティグループ対応の予備領域でなくて、複数のパリティグループに使用可能な予備領域を各ドライブに設けておき、新データ、新パリティを書き込むときに修正前のパリティグループが属するドライブ以外の２つのドライブにある２つの予備領域に

新データと新パリティを書き込んでもよい。旧データ、旧パリティは新たに予備領域とする。この方法によれば、予備領域の絶対量が少なくて済むという効果がある。

【０１２１】〔実施例５〕本発明の第５の実施例を図１７～図２０により説明する。本実施例では図１７、１８に示すように、論理グループ１０単位にサブＤＫＣ１１を設ける。そして、その内部に図１９に示すように、実施例１、２、３、４において示したキャッシュメモリ７内のアドレステーブル４０と、ＰＧ３６、サブキャッシュ３２とそれらを制御するマイクロプロセッサＭＰ３２９を持たせたものである。本実施例におけるデータの処理手順は実施例１及び２で示したものと同様である。以下には実施例１、２、３、４で示した処理手順と異なる部分のみを説明する。

【０１２２】本実施例では図１９に示すように実施例１、２、３、４で示したキャッシュメモリ７内のアドレステーブル４０を、サブＤＫＣ１１内のデータアドレステーブル（ＤＡＴ）３０に格納する。ＤＡＴ３０は格納されているテーブルの形式や機能は実施例１、２、３、４のテーブル４０と同様であるが、異なるのはデータを格納するＳＣＳＩドライブアドレスが論理グループ１０に属するものに限られている点と、書き込み、読み出しデータは保持しない専用メモリ上に構成されていることである。このため、以下では図３（ｂ）のテーブル４０をＤＡＴ３０の代りに適宜引用する。ＡＤＣ２内のＧＡＴ２３はＣＰＵ１から指示されたＣＰＵ指定アドレスから、そのＣＰＵ指定アドレスが指示する場所がＡＤＵ３内のどの論理グループ１０かを判定するのみである。

【０１２３】キャッシュメモリ７内には、その特別な領域に図２０に示すような論理グループテーブル（ＬＧＴ）６０が格納されている。ＬＧＴ６０は図２０に示すようにＣＰＵ１から指定されるＣＰＵ指定ドライブ番号４１とＣＣＨＨＲ４６に対応して、論理グループアドレス６１が決定できるテーブルとなっている。また、ＬＧＴ６０にはＣＰＵ指定アドレスに対応するデータがキャッシュメモリ７内に存在する場合、そのデータのキャッシュメモリ７内のアドレスをキャッシュアドレス４７に登録でき、また、キャッシュメモリ７内にデータが存在する場合オン（１）とし、キャッシュメモリ７内に存在しない場合オフ（０）とするキャッシュフラグ４８が用意されている。ユーザは初期設定する際に自分の使用可能な容量に対する領域を確保するが、その際にＡＤＣ２のＭＰ１２０が、ＬＧＴ６０により論理グループ１０を割当てる。

【０１２４】この時、ＭＰ２０はＬＧＴ６０にユーザが確保するために指定したＣＰＵ指定アドレスに対応する領域を登録する。そこで、実際の読みだし、書き込み処理においては、ＧＡＴ２３はＬＧＴ６０によりＣＰＵ１から指定してきたＣＰＵ指定アドレスに対応した論理グループ１０を認識することが可能となる。

【０１２５】読み出し時はＧＡＴ２３がＬＧＣにより論理グループ１０を確定し、その確定した結果をＭＰ２０に報告し、ＭＰ２０はこの当該論理グループ１０に対し読み出し要求を発行するようにＤｒｉｖｅ　ＩＦ２８に指示する。ＭＰ２０から指示を受けたＤｒｉｖｅ　ＩＦ２８は当該論理グループ１０のサブＤＫＣ１１に対し読み出し要求とＣＰＵ１が指定するＣＰＵ指定アドレスを発行する。サブＤＫＣ１１ではマイクロプロセッサ（ＭＰ）２９がこの読み出し要求のコマンドとＣＰＵ指定アドレスを受け付け、実施例１と同様Ｄｒｉｖｅ　ＩＦ２８から送られてきたＣＰＵ指定アドレスをＤＡＴ３０を参照し、当該データが格納されている論理グループ１０内の物理アドレスに変換し、この物理アドレスから当該ＳＣＳＩドライブアドレス（ＳＣＳＩドライブ番号とその中のＳＣＳＩ内Ａｄｄｒ）を確定する。このアドレスの確定後ＭＰ２９は当該ＳＣＳＩドライブ１２に対し、読み出し要求を発行する。

【０１２６】ＭＰ２９から読み出し要求を発行されたＳＣＳＩドライブ１２ではそのＳＣＳＩ内Ａｄｄｒへシーク、回転待ちを行ない、当該データの読み出しが可能になり次第、該当データをドライブアダプタ回路（Ｄｒｉｖｅ Ａｄｐ）３４に転送し、Ｄｒｉｖｅ Ａｄｐ３４はサブキャッシュメモリ３２に格納する。サブキャッシュメモリ３２に当該データの格納が完了した後、Ｄｒｉｖｅ Ａｄｐ３４はＭＰ３ ２９に格納報告を行ない、ＭＰ２９はＤＡＴ３０の当該データの当該論理アドレス４５内の当該キャッシュフラグ４８をオン（１）とする。後に当該キャッシュフラグ４８がオン（１）のデータに対し読み出しまたは書込み要求が発行された場合は、サブキャッシュ３２内で処理を行なう。

【０１２７】実施例１と同様にＭＰ２９によるＤＡＴ３０の更新が終了すると、ＭＰ２９はＡＤＣ２内のＤｒｉｖｅ　ＩＦ２８に対しデータ転送可能という応答を行ない、Ｄｒｉｖｅ　ＩＦ２８はこの応答を受け取ると、ＭＰ２０に対し報告する。ＭＰ２０はこの報告を受け取ると、キャッシュメモリ７への格納が可能なら、Ｄｒｉｖｅ　ＩＦ２８に対しサブＤＫＣ１１からデータを転送するように指示する。Ｄｒｉｖｅ　ＩＦ２８ではＭＰ２０からの指示を受けるとサブＤＫＣ１１のＭＰ２９に対し読み出し要求を発行する。この読み出し要求を受けたＭＰ２９はサブキャッシュアダプタ回路（ＳＣＡ）３１に対しサブキャッシュ３２から当該データを読みだすように指示し、ＳＣＡ３１は実際にデータを読み出してＤｒｉｖｅＩＦ２８にデータを転送する。Ｄｒｉｖｅ　ＩＦ２８がデータを受け取った後は、実施例１、２で示した処理を行なう。

【０１２８】一方、書込み時は読み出し時と同様に当該論理グループ１０を確定し、ＭＰ２０はＤｒｉｖｅ　ＩＦ２８に対し当該論理グループ１０のＭＰ２９に対し書

込み要求を発行するように指示する。当該論理グループ１０内のＭＰ３　２９は、書込み要求を受け付け、書込みデータをサブキャッシュ３２に格納した後は、図５のフローチャートに従って、実施例１、２、３、４と同様にサブＤＫＣ１１内で当該ＳＣＳＩドライブ１２群において回転位置の検出を行なう。ＭＰ３　２９は回転位置の検出を行なった後は書込みＳＣＳＩドライブ１２を確定する。ＭＰ２９により書込みＳＣＳＩドライブ１２を確定した後は実施例１、２、３、４と同様に処理する。本実施例では実施例１、２、３、４と同様の効果を実現することが可能である。

【０１２９】以上述べた各実施例においては、磁気ディスク装置を用いたシステムを例に説明したが、本発明は光ディスク装置を用いた記憶システムのような他の回転型の記録媒体に用いても同様な効果を発揮することが可能である。

【０１３０】

【発明の効果】本発明によれば、データの更新時における回転待ち時間を短縮できるため、書き込み処理を高速に行え、これにより単位時間当りのＩ／Ｏ処理件数を増加させることが可能となる。また、ＳＣＳＩドライブに障害が発生した場合、すぐに正常なＳＣＳＩドライブと交換する必要が無いため、システムを使用するユーザの負担が軽くなる。また、アドレス管理が簡単なため、ドライブの拡張も容易である。さらに、通常は使用しないスペアのＳＣＳＩドライブを、回転待ち時間の短縮という、性能向上のために使用でき、ＳＣＳＩドライブ資源の有効活用が図れる。

【図面の簡単な説明】

【図１】本発明の第１の実施例の全体構成図。

【図２】第１の実施例のクラスタ内構成図。

【図３】アドレス変換テーブルの説明図。

【図４】更新処理時のデータ移動説明図。

【図５】更新処理フローチャート。

【図６】パリティグループを構成する各データのディスク上での位置説明図。

【図７】更新処理タイミングチャート。

【図８】ディスクの回転位置検出方法説明図。

【図９】更新前後におけるパリティグループ内データの詳細説明図。

【図１０】第１の実施例の変形例になる論理グループのアドレス構成図。

【図１１】第２の実施例におけるデータ回復処理説明図。

【図１２】図１１の例の障害回復処理のフローチャート。

【図１３】第３の実施例におけるドライブ拡張例の構成図。

【図１４】図１３の例における更新処理タイミングチャート。

【図１５】ドライブ拡張時のデータ書き込み処理フローチャート。

【図１６】第４の実施例の論理グループの説明図。

【図１７】第５の実施例のシステム構成図。

【図１８】第５の実施例のクラスタ内構成図。

【図１９】第５の実施例のサブＤＫＣ内構成図。

【図２０】第５の実施例の論理グループテーブル説明図。

【図２１】従来の汎用大型計算機とＲＡＩＤ　Ｌｅｖｅ１５における更新処理説明図。

【図２２】図２１の例におけるドライブ内アドレス説明図。

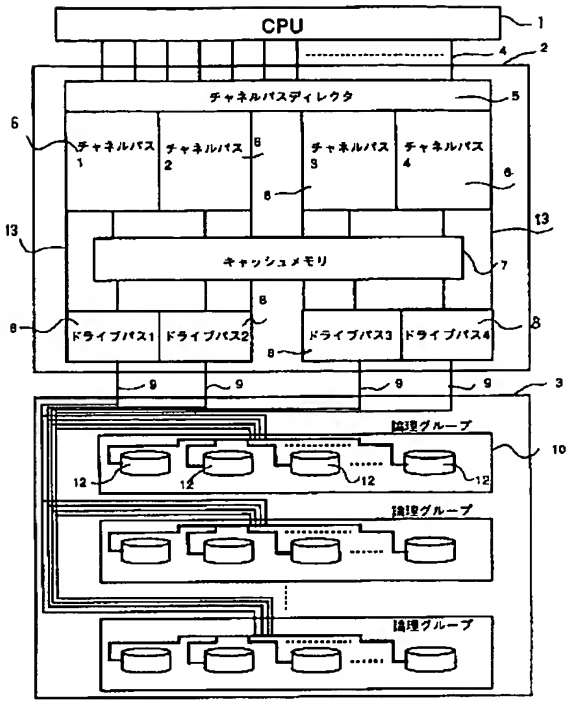
【図２３】図２１のＲＡＩＤ５における書き込み処理タイミングチャート。

【符号の説明】

１…ＣＰＵ、２…アレイディスクコントローラ（ＡＤＣ）、３…アレイディスクユニット（ＡＤＵ）、４…外部インターフェースパス、５…チャネルパスディレクタ、６…チャネルパス、７…キャッシュメモリ、８…ドライブパス、９…アレイディスクユニットパス、１０…論理グループ、１２…ＳＣＳＩドライブ、１３…クラスタ、１４…ドライブ側キャッシュアダプタ（Ｃ Ａｄｐ）、１５…インターフェースアダプタ、１６…チャネルパススイッチ、１７…制御信号線、１８…データ線、１９…パス、２０…マイクロプロセッサ１（ＭＰ１）、２１…チャネルインターフェース（ＣＨ　ＩＦ）回路、２２…データ制御回路（ＤＣＣ）、２３…グループアドレス変換回路（ＧＡＴ）、２４…チャネル側キャッシュアダプタ（Ｃ Ａｄｐ）、２８…ドライブインターフェース回路（Ｄｒｉｖｅ　ＩＦ）、２９…マイクロプロセッサ３（ＭＰ３）、３０…データアドレステーブル４０、３１…サブキャッシュアダプタ、３４…ドライブアダプタ（Ｄｒｉｖｅ Ａｄｐ）、３５…ドライブパス、３６…パリティ生成回路、４０…アドレス変換用テーブル（アドレステーブル）、４７…キャッシュアドレス、４８…キャッシュフラグ、４９…無効フラグ、６０…論理グループテーブル、６１…論理グループアドレス

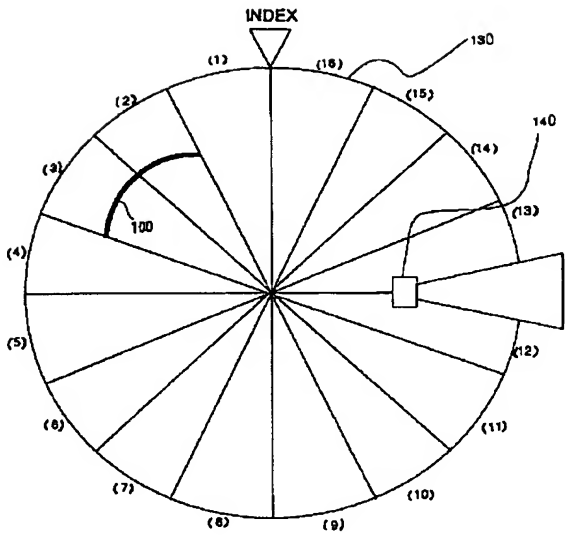
【図１】

図1



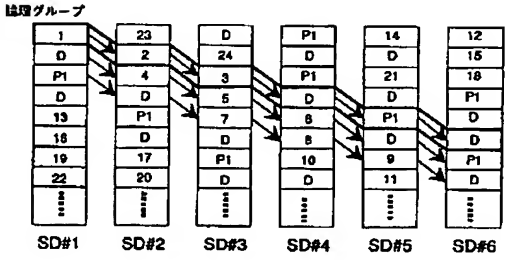
【図８】

図8



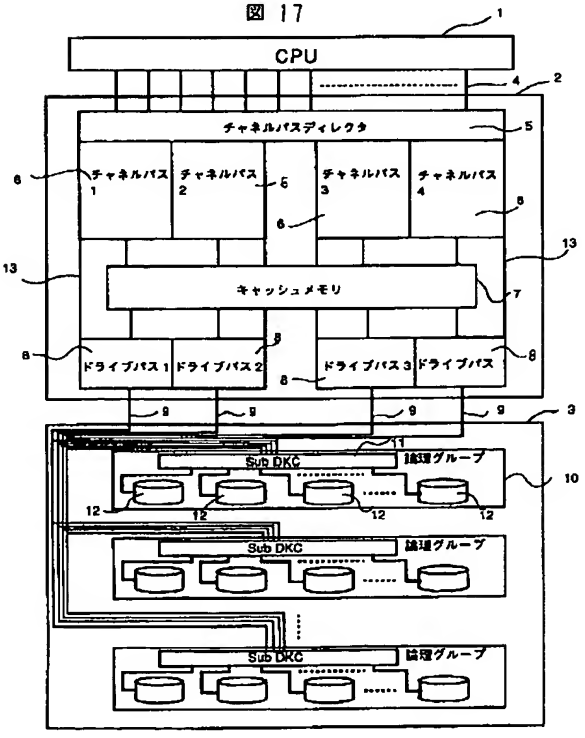
【図１０】

図 10



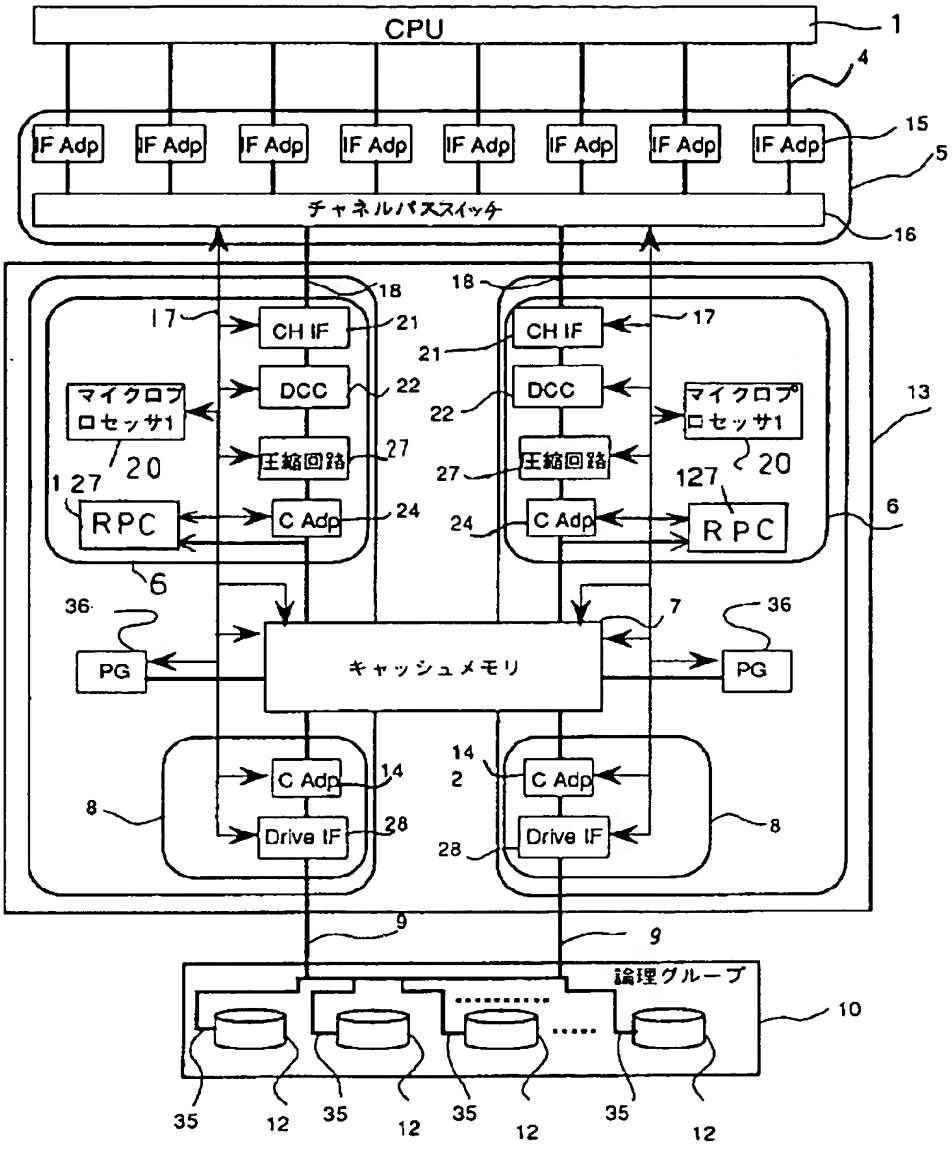
【図１７】

図 17



【図２】

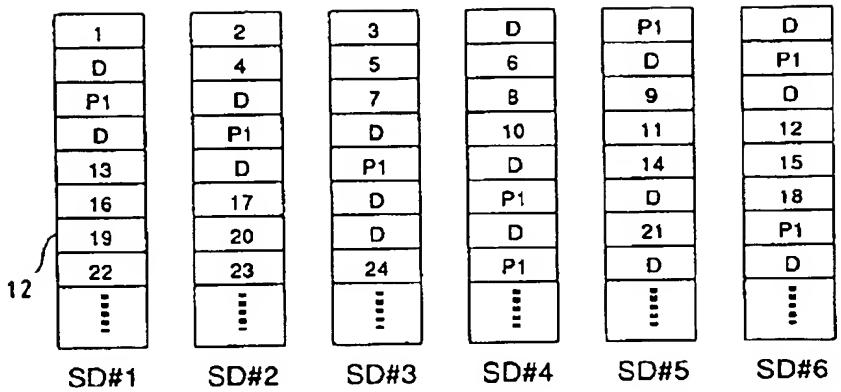
図2


CPU
1
4
IF Adp
15
5
チャネルパススイッチ
16
18
17
CH IF
21
DCC
22
マイクロプロセッサ1
20
圧縮回路
27
127
RPC
C Adp
24
6
13
36
PG
キャッシュメモリ
7
14
2
Drive IF
28
8
9
論理グループ
10
35
12

【図３】

図3

(a)



(b)

CPU指定 Drive No.	SCASI ドライブアドレス				SCSI ドライブ情報					
	SCSI 内Addr	PDrive No.	DDrive No.		SD#1	SD#2	SD#3	SD#4	SD#5	SD#6
Drive#1	DADR1	SD#5	SD#4	SD#6	Data#1	Data#2	Data#3	Dummy	Parity	Dummy
	DADR2	SD#6	SD#1	SD#5	Dummy	Data#4	Data#5	Data#6	Dummy	Parity
	DADR3	SD#1	SD#2	SD#6	Parity	Dummy	Data#7	Data#8	Data#9	Dummy
	DADR4	SD#2	SD#1	SD#3	Dummy	Parity	Dummy	Data#10	Data#11	Data#12
	DADR5	SD#3	SD#2	SD#4	Data#13	Dummy	Parity	Dummy	Data#14	Data#15
	DADR6	SD#4	SD#3	SD#5	Data#16	Data#17	Dummy	Parity	Dummy	Data#18
	DADR7	SD#6	SD#3	SD#4	Data#19	Data#20	Dummy	Dummy	Data#21	Parity
	DADR8	SD#4	SD#5	SD#6	Data#22	Data#23	Data#24	Parity	Dummy	Dummy

41 42 43 40 44 50 51 45 46 47 48 49

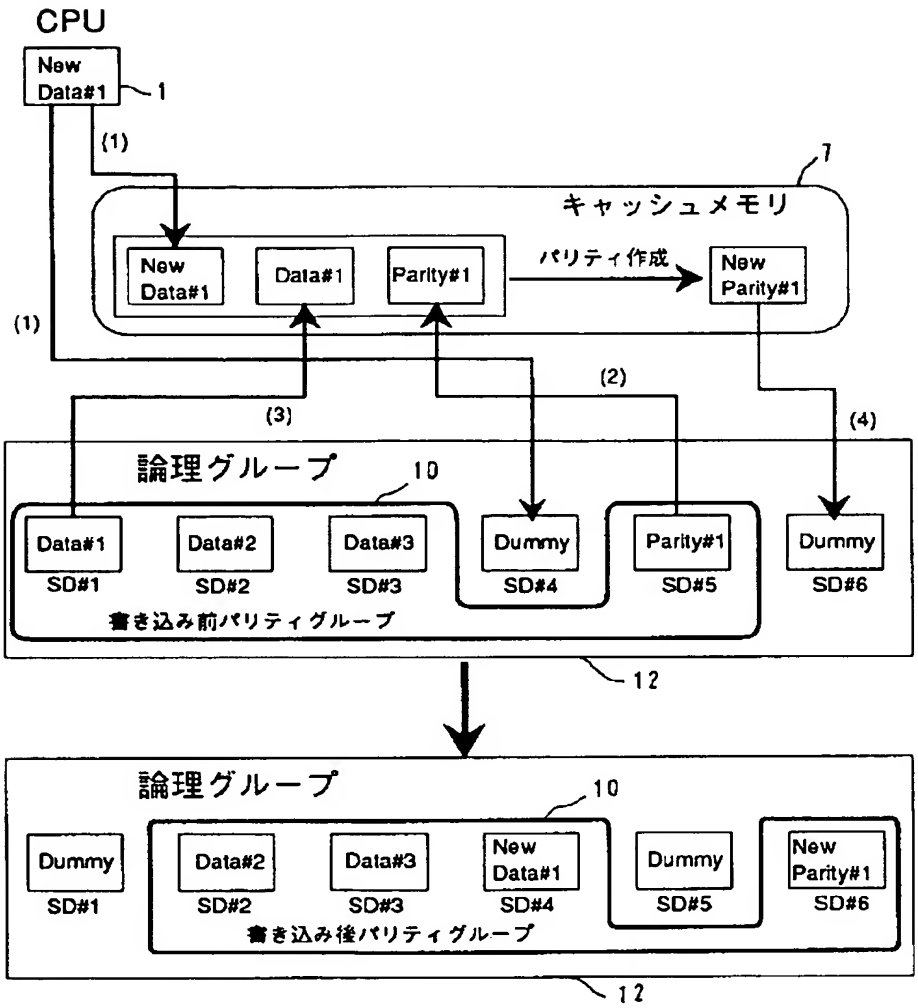
PDrive No.:パリティドライブ番号

DDrive No.:ダミードライブ番号

CCHHR	キャッシュアドレス	Cache Flag	無効 Flag
ADR 8	C ADR2,1	0	0

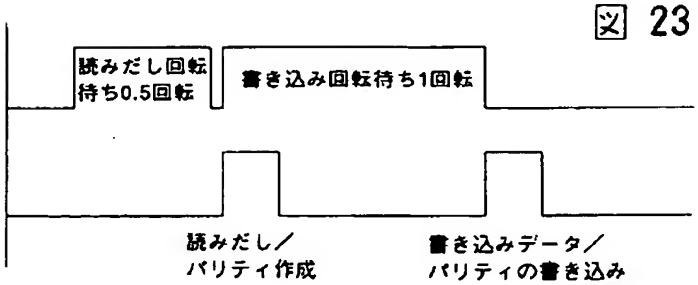
【図４】

図4


CPU
New Data#1
1
(1)
7
キャッシュメモリ
New Data#1
Data#1
Parity#1
パリティ作成
New Parity#1
(1)
(2)
(3)
(4)
論理グループ
10
Data#1
SD#1
Data#2
SD#2
Data#3
SD#3
Dummy
SD#4
Parity#1
SD#5
Dummy
SD#6
書き込み前パリティグループ
12
論理グループ
10
Dummy
SD#1
Data#2
SD#2
Data#3
SD#3
New Data#1
SD#4
Dummy
SD#5
New Parity#1
SD#6
書き込み後パリティグループ
12

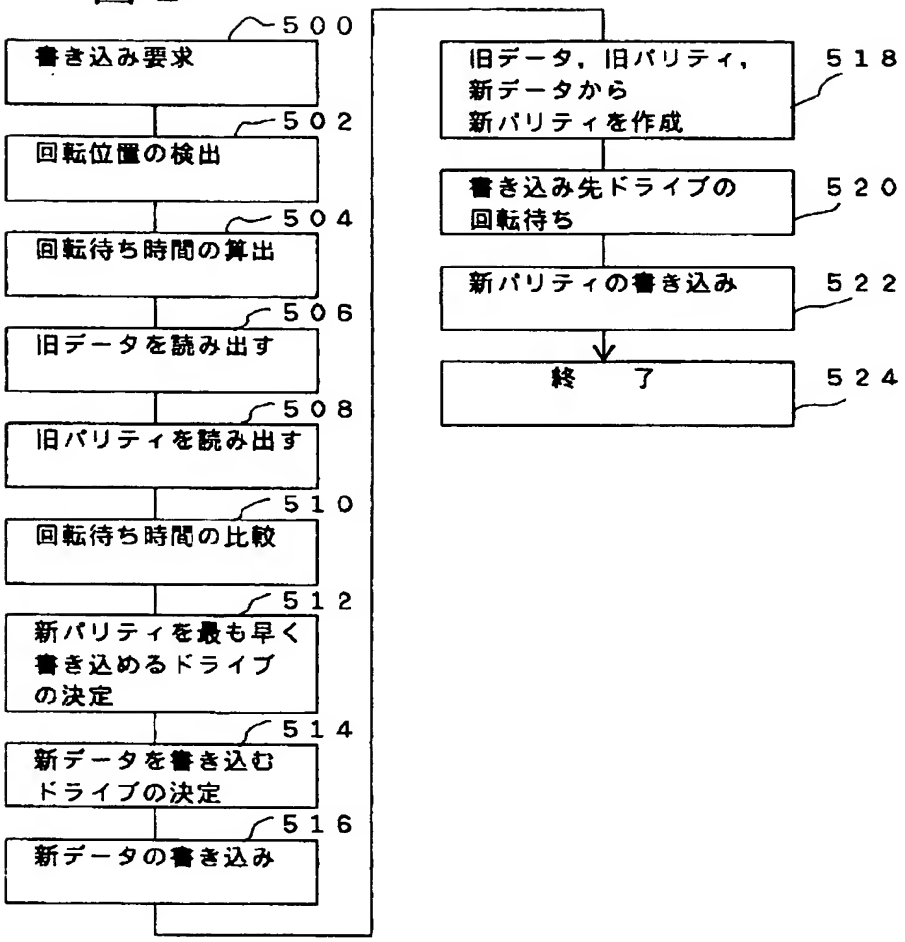
【図２３】

図 23


読みだし回転待ち0.5回転
書き込み回転待ち1回転
読みだし／パリティ作成
書き込みデータ／パリティの書き込み

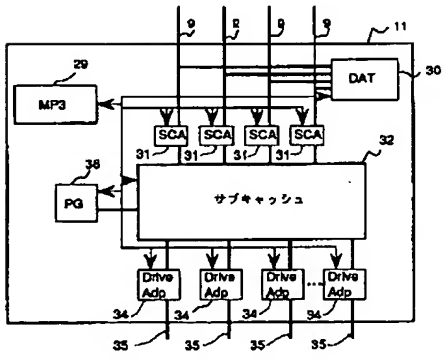
【図５】

図 5



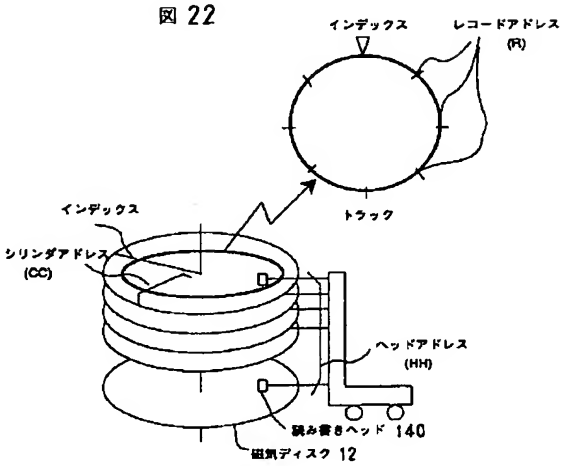
【図１９】

図 19



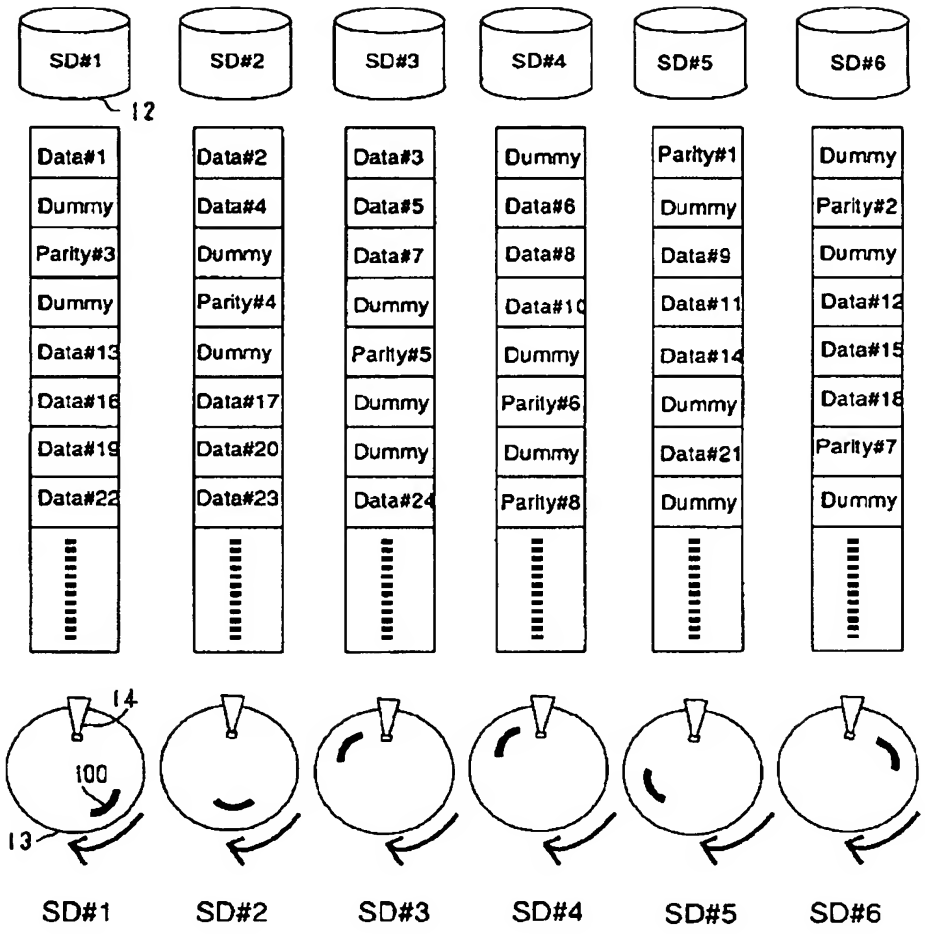
【図２２】

図 22



【図６】

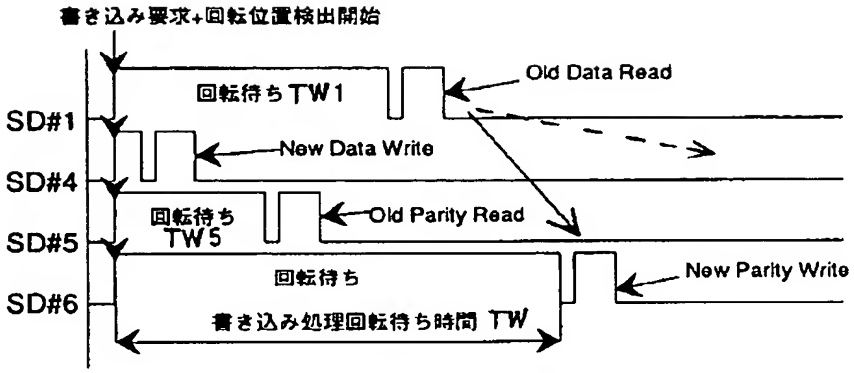
図6


SD#1
SD#2
SD#3
SD#4
SD#5
SD#6
12
Data#1
Dummy
Parity#3
Dummy
Data#13
Data#16
Data#19
Data#22
Data#2
Data#4
Dummy
Parity#4
Dummy
Data#17
Data#20
Data#23
Data#3
Data#5
Data#7
Dummy
Parity#5
Dummy
Dummy
Data#24
Dummy
Data#6
Data#8
Data#10
Dummy
Parity#6
Dummy
Parity#8
Parity#1
Dummy
Data#9
Data#11
Data#14
Dummy
Data#21
Dummy
Dummy
Parity#2
Dummy
Data#12
Data#15
Data#18
Parity#7
Dummy
14
100
13
SD#1
SD#2
SD#3
SD#4
SD#5
SD#6

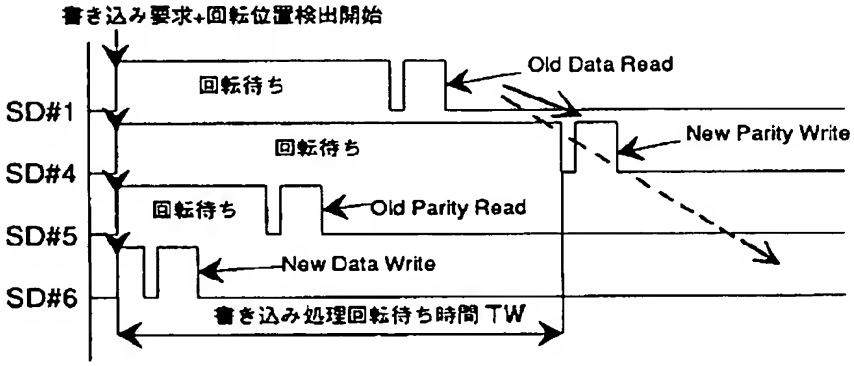
【図７】

図7

(a)

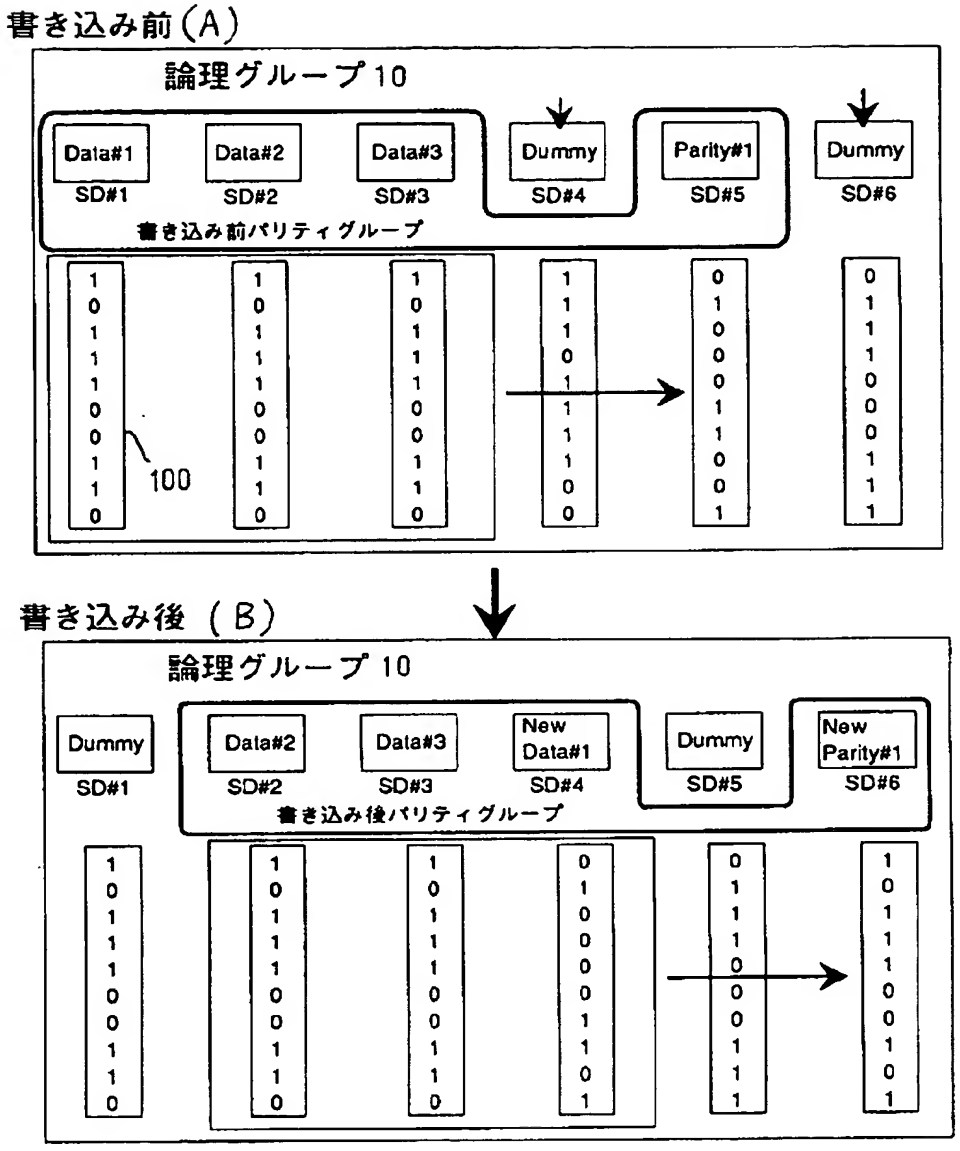

書き込み要求+回転位置検出開始
回転待ちTW1
Old Data Read
SD#1
New Data Write
SD#4
回転待ち
TW5
Old Parity Read
SD#5
回転待ち
New Parity Write
SD#6
書き込み処理回転待ち時間 TW

(b)


書き込み要求+回転位置検出開始
回転待ち
Old Data Read
SD#1
回転待ち
New Parity Write
SD#4
回転待ち
Old Parity Read
SD#5
New Data Write
SD#6
書き込み処理回転待ち時間 TW

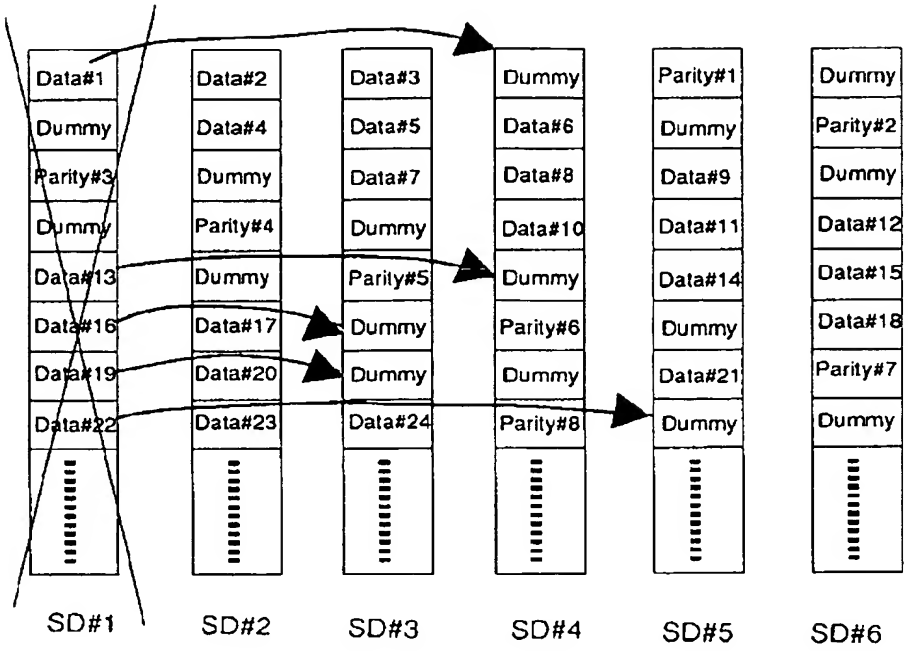
【図９】

図9


書き込み前（A）
論理グループ10
Data#1
SD#1
Data#2
SD#2
Data#3
SD#3
Dummy
SD#4
Parity#1
SD#5
Dummy
SD#6
書き込み前パリティグループ
100
書き込み後（B）
論理グループ10
Dummy
SD#1
Data#2
SD#2
Data#3
SD#3
New Data#1
SD#4
Dummy
SD#5
New Parity#1
SD#6
書き込み後パリティグループ

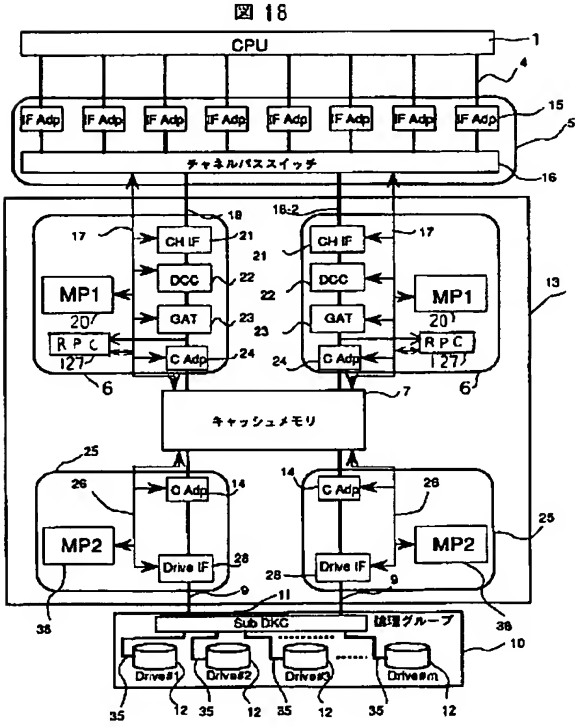
【図１１】

図 11



【図１８】

図 18



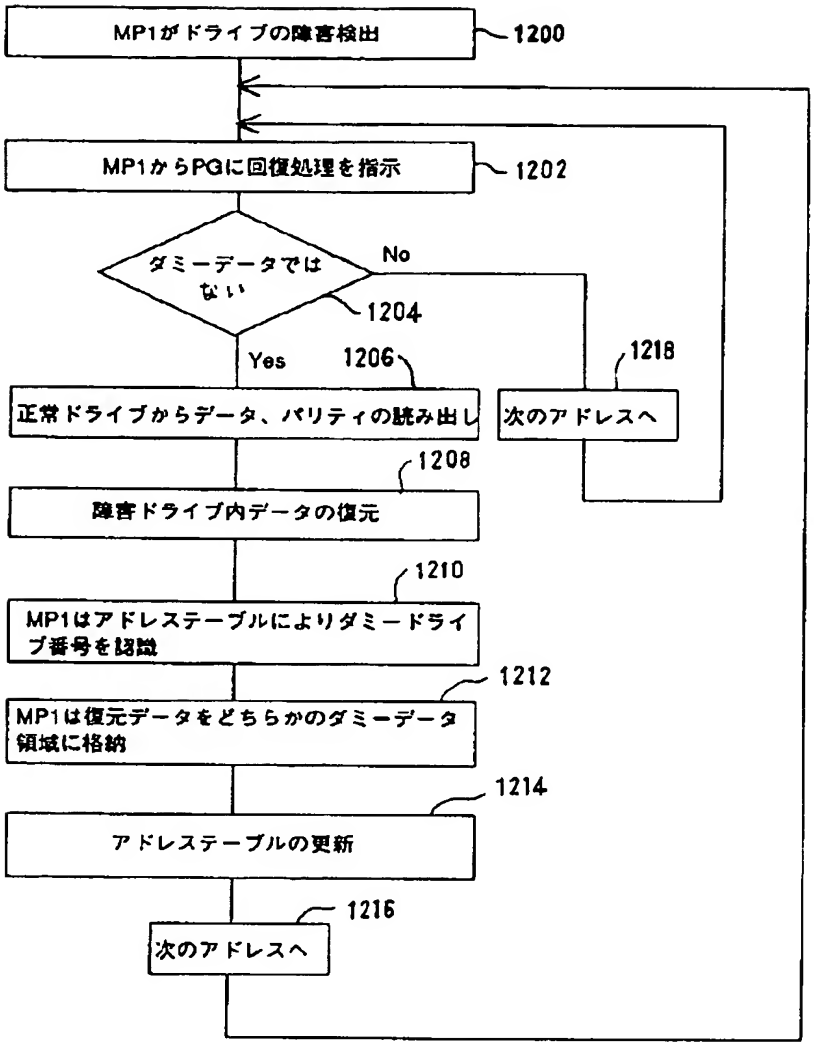
【図２０】

図 20

CPU指定アドレス				
CPU指定 Drive No.	CCHHR	論理グループアドレス	キャッシュアドレス	Cache Flag
Drive#1	ADR 1	LADR 1	——	——
	ADR 2	LADR 3	——	——
	ADR 3	LADR 8	——	——
	⋮	⋮	⋮	⋮
Drive#2	ADR 1	LADR 2	——	0
	ADR 2	LADR 1	CADR1,5	1
	ADR 3	LADR 5	CADR1,8	1
	ADR 4	LADR 4	CADR1,6	1
⋮	⋮	⋮	⋮	⋮

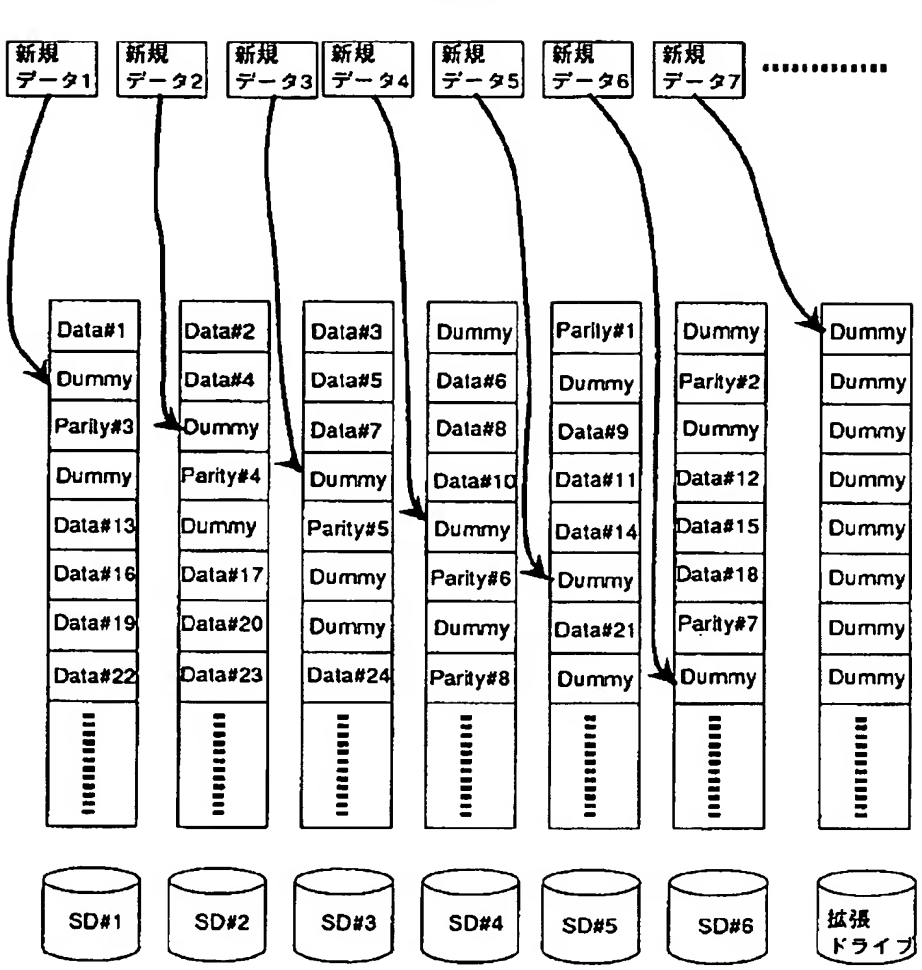
【図１２】

図 12


MP1がドライブの障害検出
1200
MP1からPGに回復処理を指示
1202
ダミーデータではない
No
1204
Yes
1206
正常ドライブからデータ、パリティの読み出し
1218
次のアドレスへ
1208
障害ドライブ内データの復元
1210
MP1はアドレステーブルによりダミードライブ番号を認識
1212
MP1は復元データをどちらかのダミーデータ領域に格納
1214
アドレステーブルの更新
1216
次のアドレスへ

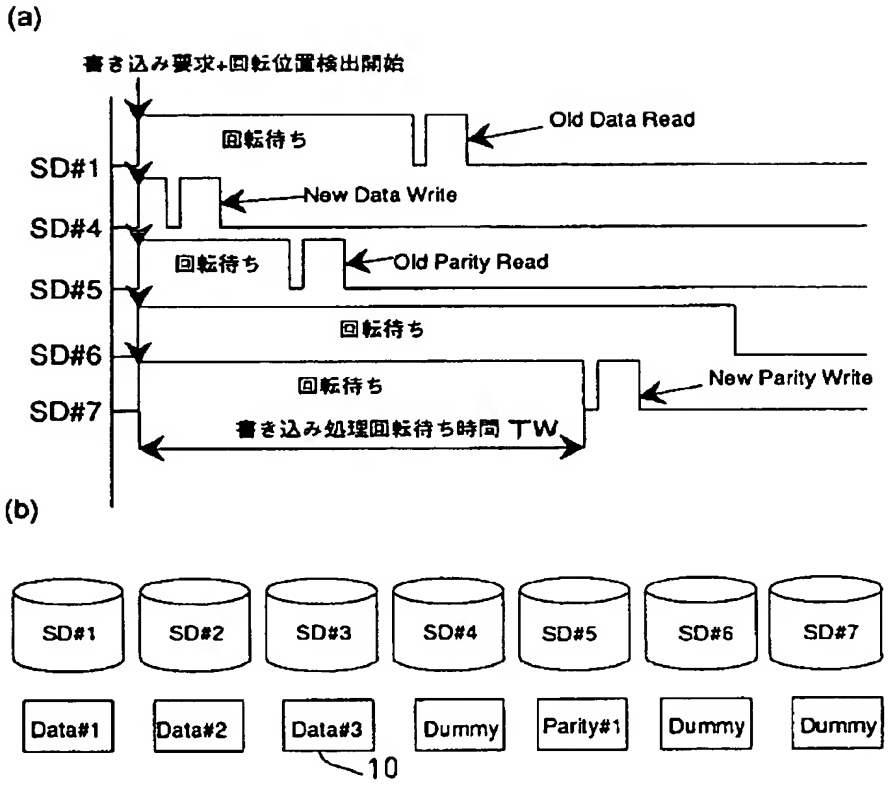
【図１３】

図 13


CPU
新規データ1
新規データ2
新規データ3
新規データ4
新規データ5
新規データ6
新規データ7
Data#1
Dummy
Parity#3
Dummy
Data#13
Data#16
Data#19
Data#22
Data#2
Data#4
Dummy
Parity#4
Dummy
Data#17
Data#20
Data#23
Data#3
Data#5
Data#7
Dummy
Parity#5
Dummy
Dummy
Data#24
Dummy
Data#6
Data#8
Data#10
Dummy
Parity#6
Dummy
Parity#8
Parity#1
Dummy
Data#9
Data#11
Data#14
Dummy
Data#21
Dummy
Dummy
Parity#2
Dummy
Data#12
Data#15
Data#18
Parity#7
Dummy
Dummy
Dummy
Dummy
Dummy
Dummy
Dummy
Dummy
Dummy
Dummy
SD#1
SD#2
SD#3
SD#4
SD#5
SD#6
拡張ドライブ

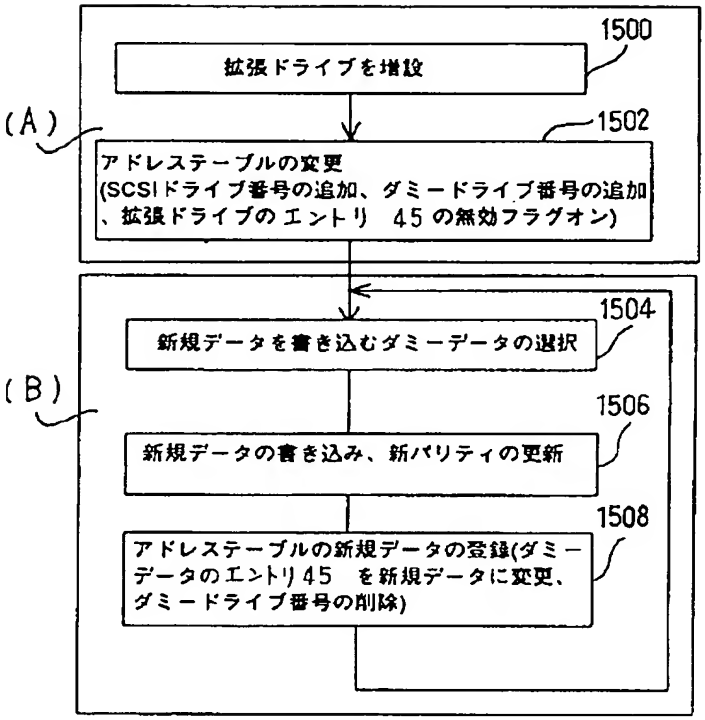
【図１４】

図 14


(a)
書き込み要求+回転位置検出開始
回転待ち
Old Data Read
SD#1
New Data Write
SD#4
回転待ち
Old Parity Read
SD#5
回転待ち
SD#6
回転待ち
New Parity Write
SD#7
書き込み処理回転待ち時間 TW
(b)
SD#1
SD#2
SD#3
SD#4
SD#5
SD#6
SD#7
Data#1
Data#2
Data#3
Dummy
Parity#1
Dummy
Dummy
10

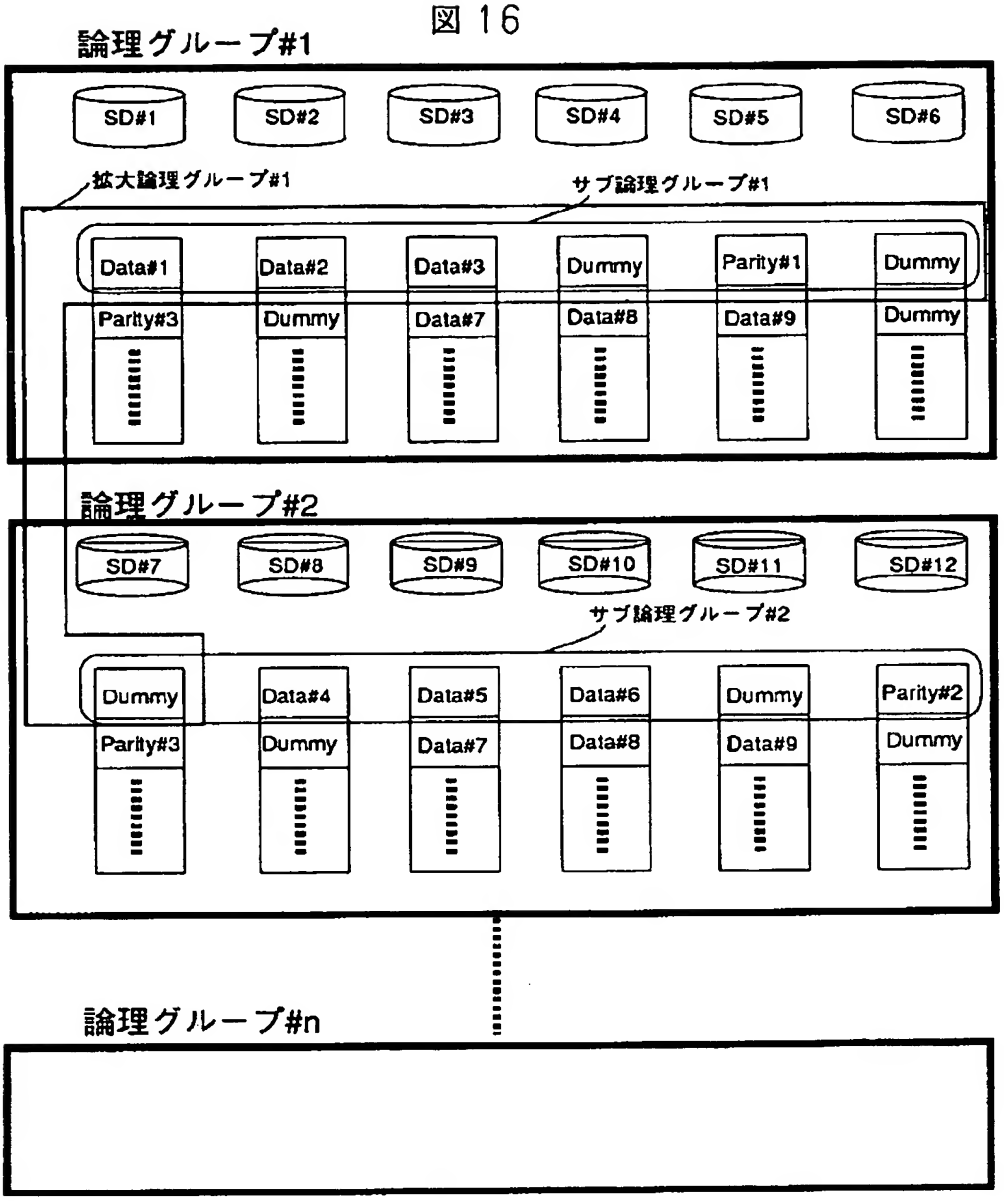
【図１５】

図 15


1500
拡張ドライブを増設
(A)
1502
アドレステーブルの変更
(SCSIドライブ番号の追加、ダミードライブ番号の追加
、拡張ドライブのエントリ 45 の無効フラグオン)
1504
新規データを書き込むダミーデータの選択
(B)
1506
新規データの書き込み、新パリティの更新
1508
アドレステーブルの新規データの登録(ダミー
データのエントリ45 を新規データに変更、
ダミードライブ番号の削除)

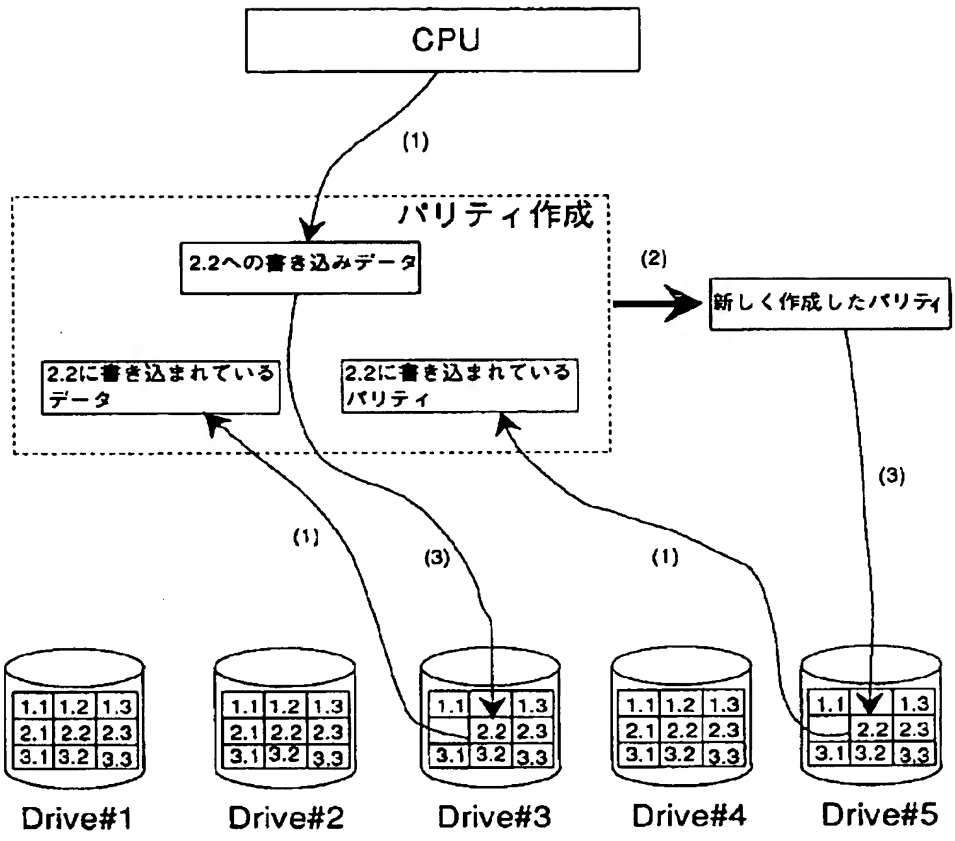
【図１６】

図 16


論理グループ#1
SD#1
SD#2
SD#3
SD#4
SD#5
SD#6
拡大論理グループ#1
サブ論理グループ#1
Data#1
Data#2
Data#3
Dummy
Parity#1
Dummy
Parity#3
Dummy
Data#7
Data#8
Data#9
Dummy
論理グループ#2
SD#7
SD#8
SD#9
SD#10
SD#11
SD#12
サブ論理グループ#2
Dummy
Data#4
Data#5
Data#6
Dummy
Parity#2
Parity#3
Dummy
Data#7
Data#8
Data#9
Dummy
論理グループ#n

【図２１】

図 21


CPU
(1)
パリティ作成
2.2への書き込みデータ
(2)
新しく作成したパリティ
2.2に書き込まれているデータ
2.2に書き込まれているパリティ
(3)
(1)
(3)
(1)
1.1 1.2 1.3
2.1 2.2 2.3
3.1 3.2 3.3
Drive#1
Drive#2
Drive#3
Drive#4
Drive#5